

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: July 14, 2000, 09:33:58 ; Search time 31.39 Seconds
(without alignments)
25.226 Million cell updates/sec

Title: ALPHA-CHAIN
Perfect score: 147
Sequence: 1 apdvdcptclqenpfqpgapil 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 83857 seqs, 30454973 residues

Total number of hits satisfying chosen parameters: 83856

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 1000 summaries

Database : SwissProt_38.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95	64.6	120	1	GLHA_MACMU
2	91	61.9	120	1	GLHA_MELGA
3	90	61.2	96	1	GLHA_STRCA
4	85	57.8	120	1	GLHA_BOVIN
5	85	57.8	120	1	GLHA_SHEEP
6	84	57.1	120	1	GLH1_RAT
7	84	57.1	120	1	GLH2_RAT
8	84	57.1	120	1	GLHA_MOUSE
9	82	55.8	96	1	GLHA_RABIT
10	82	55.8	120	1	GLHA_MACRU
11	82	55.8	120	1	GLHA_PIG
12	80	54.4	120	1	GLHA_POUAS
13	78	53.1	120	1	GLHA_CALVA
14	78	53.1	120	1	GLHA_HORSE
15	77	52.4	96	1	GLHA_PHYCA
16	74	50.3	96	1	GLHA_BALAC
17	73	49.7	116	1	GLHA_CLAGA
18	73	49.7	117	1	GLHA_ANGAN
19	73	49.7	118	1	GLH1_CYPCA
20	73	49.7	118	1	GLH2_CYPCA
21	73	49.7	118	1	GLHA_CTEID
22	73	49.7	118	1	GLHA_HYPMO
23	72	49.0	93	1	GLHA_MURCI
24	72	49.0	108	1	GLH1_ONCKE
25	53	36.1	97	1	GLHA_RANCA
26	52	35.4	1522	1	MRP3_RAT
27	50.5	34.4	522	1	FINC_CANFA
28	50.5	34.4	522	1	FINC_HORSE
29	50.5	34.4	2265	1	FINC_BOVIN
30	50.5	34.4	2477	1	FINC_MOUSE
31	50	34.0	262	1	PROF_ECOLI
32	49.5	33.7	2386	1	FINC_HUMAN
33	49.5	33.7	2477	1	FINC_RAT

34	49	33.3	278	1	HUPJ_RHOCA
35	48	32.7	117	1	GLHA_MORSA
36	48	32.7	1420	1	SRB9_YEAST
37	47	32.0	94	1	GLHA_THOGB
38	47	32.0	122	1	RPB9_YEAST
39	46.5	31.6	260	1	HXC9_MOUSE
40	46	31.3	288	1	MOV9_PSVJ
41	45.5	31.0	1034	1	GCSP_FLAAPR
42	45.5	31.0	1034	1	GCSP_FLAAN
43	45.5	31.0	1034	1	GCSP_FLAATR
44	45.5	31.0	1037	1	GCSA_FLAAPR
45	45	30.6	157	1	VE5_RHPV1
46	45	30.6	529	1	PUR2_ARATH
47	44.5	30.3	360	1	CCD2_CAEEL
48	44.5	30.3	397	1	CCST_BOVIN
49	44	29.9	367	1	FLGI_PSEPU
50	44	29.9	577	1	YQ4B_CAEEL
51	44	29.9	585	1	CALI_HUMAN
52	44	29.9	588	1	CALI_BOVIN
53	44	29.9	1640	1	CLH2_HUMAN
54	44	29.9	1711	1	PTPO_RAT
55	44	29.9	4543	1	LRP1_CHICK
56	44	29.9	4544	1	LRP1_HUMAN
57	43.5	29.6	402	1	GUN1_HUMGR
58	43.5	29.6	435	1	RAG2_XENLA
59	43.5	29.6	520	1	UBA2_YEAST
60	43.5	29.6	636	1	BMPI_XENLA
61	43.5	29.6	707	1	TSHE_SALSA
62	43	29.3	139	1	TSHE_ONCMY
63	43	29.3	147	1	MDHM_BRANA
64	43	29.3	341	1	Z185_MOUSE
65	43	29.3	352	1	SYA_PIRHO
66	43	29.3	915	1	FLGI_CAUCR
67	42.5	28.9	370	1	SP97_HUMAN
68	42.5	28.9	904	1	GLH2_ONCKE
69	42	28.6	114	1	GLHA_ACALA
70	42	28.6	117	1	YP96_CAEEL
71	42	28.6	256	1	MDHM_CITLA
72	42	28.6	347	1	RADA_HAEN
73	42	28.6	458	1	PEPS_ASPSA
74	42	28.6	523	1	AF17_HUMAN
75	42	28.6	1093	1	IOVO_FRAFR
76	41.5	28.2	53	1	IOVO_BAMTH
77	41.5	28.2	56	1	IOBA_CHICK
78	41.5	28.2	424	1	GAE_HUMAN
79	41.5	28.2	506	1	UROT_RAT
80	41.5	28.2	559	1	MYPC_CHICK
81	41.5	28.2	1271	1	PYRL_DROME
82	41.5	28.2	2236	1	AHNR_HUMAN
83	41.5	28.2	2960	1	PHNA_STRMU
84	41	27.9	33	1	RL32_DBTIA
85	41	27.9	60	1	TSHE_HUMAN
86	41	27.9	138	1	YON5_CAEEL
87	41	27.9	253	1	ACTR_MOUSE
88	41	27.9	296	1	ACTR_MOUSE
89	41	27.9	297	1	CC34_CAEEL
90	41	27.9	298	1	RM03_HUMAN
91	41	27.9	348	1	MPK6_ARATH
92	41	27.9	395	1	SMO_HUMAN
93	41	27.9	787	1	YNN2_YEAST
94	41	27.9	1056	1	LACC_STAAR
95	40.5	27.6	310	1	CGVK_HSVSA
96	40.5	27.6	431	1	EGFR_HUMAN
97	40.5	27.6	1210	1	MYPC_HUMAN
98	40.5	27.6	1274	1	NME4_MOUSE
99	40.5	27.6	1323	1	RRPO_P1AMV
100	40.5	27.6	1385	1	PEP1_YEAST
101	40.5	27.6	1579	1	CHD4_HUMAN
102	40.5	27.6	1912	1	RRPL_RVFPVZ
103	40.5	27.6	2149	1	HEAL_ANAVA
104	40	27.2	109	1	LSHB_RANCA
105	40	27.2	112	1	RPB9_DROME
106	40	27.2	129	1	

Q03009	rhodobacter
Q91119	morone saxa
P38931	saccharomyc
P37204	thunnus obe
P27999	saccharomyc
P09633	mus musculus
P22117	peanut stun
P49362	flaveria pr
O49850	flaveria an
O49852	flaveria tr
P49361	flaveria pr
P24834	rhesus papi
P52420	arabidopsis
P35799	caenorhabdi
P25285	bos taurus
Q52082	pseudomonas
Q17426	caenorhabdi
Q13939	homo sapien
Q28068	bos taurus
P36715	homo sapien
Q4612	rattus norv
P98157	gallus gall
Q07954	homo sapien
P56680	humicola in
Q12622	humicola gr
Q91830	xenopus lae
P52488	saccharomyc
P98070	xenopus lae
O73824	salmo salar
P37240	oncorhynch
Q43744	brassica ra
Q82394	mus musculu
Q58035	pyrococcus
P33979	caulobacter
Q12959	homo sapien
P13153	oncorhynch
P45266	haemophilus
P17783	citruslul 1
P52719	aspergillus
P55198	homo sapien
P05597	francolinus
P52259	bambusicola
P27092	gallus gall
P78334	homo sapien
P19637	rattus norv
Q90688	gallus gall
P05990	drosophila
Q09666	homo sapien
Q02419	streptococc
P49228	deinococcus
P01222	homo sapien
Q09526	caenorhabdi
Q84326	mus musculu
P70115	mesocricetu
P34687	caenorhabdi
P09001	homo sapien
Q39026	arabidopsis
Q99835	homo sapien
P53914	saccharomyc
P11099	staphylococ
Q01015	herpesvirus
P00533	homo sapien
Q14896	homo sapien
Q03331	mus musculu
Q07518	plantago as
P32319	saccharomyc
Q14839	homo sapien
P27316	rift valley
P46049	anabaena va
P80071	rana catesb
P36958	drosophila

107	40	27.2	261	1	HIS4_YEAST	P40545	saccharomyc	180	39	26.5	1021	1	BUB1_YEAST	P41695	saccharomyc
108	40	27.2	266	1	HES4_ANASP	P18500	anabaena sp	181	39	26.5	1039	1	YR71_CAEL	Q09564	caenorhabdi
109	40	27.2	394	1	IPOU_DROME	P24350	drosophila	182	39	26.5	1057	1	POL_SIVAI	Q02836	simian immu
110	40	27.2	539	1	L114_CAEL	P24350	drosophila	183	39	26.5	1064	1	Y108_METU	Q60307	methanococ
111	40	27.2	556	1	NUMB_DROME	P16554	drosophila	184	39	26.5	1131	1	PHYA_MAIZE	P19862	zea mays (m
112	40	27.2	572	1	2YX_HUMAN	Q15942	homo sapien	185	39	26.5	1222	1	YMH5_CAEL	P34472	caenorhabdi
113	40	27.2	574	1	5NTD_BOVIN	Q05977	bos taurus	186	39	26.5	1435	1	EBA1_PLAPC	P19214	plasmodium
114	40	27.2	574	1	5NTD_HUMAN	P21589	homo sapien	187	39	26.5	1511	1	PDH5_YEAST	P33302	saccharomyc
115	40	27.2	576	1	5NTD_RAT	P21588	rattus norv	188	39	26.5	1527	1	MRP3_HUMAN	O15438	homo sapien
116	40	27.2	671	1	KPC1_HUMAN	P05771	homo sapien	189	39	26.5	1884	1	RRP5_HUMAN	Q14690	homo sapien
117	40	27.2	671	1	KPC1_RABIT	P05772	oryctolagus	190	39	26.5	3084	1	LMAL_MOUSE	P19137	mus musculus
118	40	27.2	671	1	KPC1_RAT	P04410	rattus norv	191	38.5	26.2	54	1	IOVO_MEGFR	P05580	megapodius
119	40	27.2	673	1	KPC2_BOVIN	P05126	bos taurus	192	38.5	26.2	56	1	IOVO_FRAPO	P05598	francolinus
120	40	27.2	673	1	KPC2_HUMAN	P05127	homo sapien	193	38.5	26.2	207	1	Y126_MYCTU	Q10972	mycobacteri
121	40	27.2	673	1	KPC2_MOUSE	P04411	mus musculus	194	38.5	26.2	350	1	XYLA_PSEPD	P21394	pseudomonas
122	40	27.2	673	1	KPC3_RABIT	P05773	oryctolagus	195	38.5	26.2	353	1	OM51_HAEIN	P43840	haemophilus
123	40	27.2	678	1	NCPR_SCHPO	P36587	schizosacch	196	38.5	26.2	398	1	SNF_SYNY3	P73345	synecocyst
124	40	27.2	881	1	Y310_HUMAN	Q15027	homo sapien	197	38.5	26.2	426	1	ICSB_HUMAN	Q02556	homo sapien
125	40	27.2	957	1	NIRB_KLEPN	Q06458	klebsiella	198	38.5	26.2	532	1	CPP1_DROAC	O44220	drosophila
126	40	27.2	1145	1	POL_EIAVY	P03371	equine infe	199	38.5	26.2	789	1	CAD6_RAT	P55280	rattus norv
127	40	27.2	1146	1	POL_EIAV9	P11204	equine infe	200	38.5	26.2	790	1	CAD6_HUMAN	P55285	homo sapien
128	40	27.2	1146	1	POL_EIAVC	P32542	equine infe	201	38.5	26.2	790	1	CAD6_MOUSE	P97326	mus musculus
129	40	27.2	1849	1	IGM4_HAEIN	P45386	haemophilus	202	38.5	26.2	835	1	TF1B_HUMAN	Q13263	homo sapien
130	40	27.2	4303	1	PKD1_HUMAN	P98161	homo sapien	203	38.5	26.2	911	1	SP97_RAT	Q62696	rattus norv
131	39.5	26.9	248	1	SOD1_PLEBO	P50058	plectonema	204	38.5	26.2	946	1	CJTC_YEAST	P07245	s c-1-tetra
132	39.5	26.9	249	1	GRPE_SYNY3	Q59978	synecocyst	205	38.5	26.2	1396	1	VCAP_VZVD	P09245	varicella-z
133	39.5	26.9	259	1	HXC9_FUGRU	Q42502	fugu rubrip	206	38.5	26.2	1487	1	ICP4_HSVBK	P28925	equine herp
134	39.5	26.9	262	1	END8_ECOLI	P05405	escherichia	207	38.5	26.2	1487	1	ICP4_HSVBK	P17473	equine herp
135	39.5	26.9	376	1	FMOD_HUMAN	Q06828	homo sapien	208	38.5	26.2	1690	1	CA44_HUMAN	P53420	homo sapien
136	39.5	26.9	449	1	PCO1_HUMAN	Q15113	homo sapien	209	38.5	26.2	2056	1	CBP1_CAEL	P34545	caenorhabdi
137	39.5	26.9	815	1	MK07_HUMAN	Q13164	homo sapien	210	38.5	26.2	2769	1	THVG_BOVIN	P01267	bos taurus
138	39.5	26.9	937	1	SYL_METH	Q27552	methanobact	211	38	25.9	103	1	YGU3_YEAST	P50397	saccharomyc
139	39.5	26.9	1246	1	SKIW_HUMAN	Q15477	homo sapien	212	38	25.9	261	1	Y4ZC_RHISN	P55730	rhizobium s
140	39.5	26.9	1270	1	MYPC_MOUSE	O70468	mus musculus	213	38	25.9	272	1	TCIF_HUMAN	Q15583	homo sapien
141	39	26.5	111	1	PSSP_PIG	Q02826	sus scrofa	214	38	25.9	276	1	MYO1_ONCMY	Q91205	oncorhynch
142	39	26.5	191	1	Y335_MYCGE	P47577	mycoplasma	215	38	25.9	327	1	FASA_MOUSE	P25446	mus musculus
143	39	26.5	212	1	XCDC_ECOLI	P75899	escherichia	216	38	25.9	350	1	GLCE_ECOLI	P52073	escherichia
144	39	26.5	285	1	HEA2_ANAVA	P46048	anabaena va	217	38	25.9	360	1	PIW7_HUMAN	P35236	homo sapien
145	39	26.5	285	1	DDH1_HUMAN	O94760	homo sapien	218	38	25.9	364	1	ARRA_DROME	P15372	drosophila
146	39	26.5	285	1	DDH1_RAT	O08557	rattus norv	219	38	25.9	375	1	NPL4_HUMAN	Q99733	homo sapien
147	39	26.5	297	1	ACTR_HUMAN	Q01718	homo sapien	220	38	25.9	384	1	SSR4_MOUSE	P49660	mus musculus
148	39	26.5	297	1	APOH_RAT	P26644	rattus norv	221	38	25.9	413	1	GAT1_MOUSE	P17679	mus musculus
149	39	26.5	332	1	FASA_PIG	O77736	sus scrofa	222	38	25.9	413	1	GAT1_RAT	P43429	rattus norv
150	39	26.5	338	1	GLB_ASCSU	P28316	ascaris suu	223	38	25.9	413	1	NUOD_RHOCA	O07310	rhodobacter
151	39	26.5	345	1	APOH_BOVIN	P17690	bos taurus	224	38	25.9	413	1	YWS2_CAEL	Q10938	caenorhabdi
152	39	26.5	373	1	CD14_BOVIN	Q95122	bos taurus	225	38	25.9	422	1	CG2A_SPI50	O49462	spisula sol
153	39	26.5	397	1	YMB1_CAEL	Q03599	caenorhabdi	226	38	25.9	437	1	PKNA_MYCLE	P54743	mycobacteri
154	39	26.5	416	1	NGFR_CHICK	P18519	gallus gall	227	38	25.9	450	1	YOE3_CAEL	P54743	caenorhabdi
155	39	26.5	417	1	AIAB_CANFA	P11615	canis famil	228	38	25.9	453	1	A2AB_RAT	P19328	rattus norv
156	39	26.5	427	1	YEBS_ECOLI	P76271	escherichia	229	38	25.9	467	1	CBPA_DICDI	P35085	dictyostell
157	39	26.5	442	1	BNE_DROME	P29746	drosophila	230	38	25.9	476	1	THB2_HUMAN	P37243	homo sapien
158	39	26.5	452	1	2185_HUMAN	Q15231	homo sapien	231	38	25.9	489	1	ANSP_STRCO	Q9x7p0	streptomyce
159	39	26.5	466	1	VL2_HPV31	P17389	human papil	232	38	25.9	498	1	IAP2_DROME	Q24307	drosophila
160	39	26.5	490	1	FAIO_RABIT	O19045	oryctolagus	233	38	25.9	509	1	2KRL_CHICK	Q30373	gallus gall
161	39	26.5	514	1	AIAB_MOUSE	P97717	mus musculus	234	38	25.9	542	1	GUAA_SYNY3	P49057	synecocyst
162	39	26.5	515	1	AIAB_MESAU	P18841	mesocricetu	235	38	25.9	586	1	VDMB_BPT7	P03694	bacterioph
163	39	26.5	515	1	AIAB_RAT	P15823	rattus norv	236	38	25.9	596	1	PFOB_LYCFS	Q08304	lycopersico
164	39	26.5	519	1	AIAB_HUMAN	P35368	homo sapien	237	38	25.9	597	1	DYI3_ANTCR	Q16960	anthocidari
165	39	26.5	524	1	CP51_UCNGE	O14442	uncinula ne	238	38	25.9	598	1	NR42_MOUSE	Q06219	mus musculus
166	39	26.5	537	1	MYPH_CHICK	Q05623	gallus gall	239	38	25.9	602	1	TRAN_ECOLI	P24082	escherichia
167	39	26.5	564	1	2YX_MOUSE	Q62523	mus musculus	240	38	25.9	615	1	ENV_J5RV	P31621	sheep pulmo
168	39	26.5	576	1	YITO_YEAST	P40568	saccharomyc	241	38	25.9	646	1	SKB1_SCHPO	P78963	schizosacch
169	39	26.5	587	1	COAT_PAVL3	P36310	parvovirus	242	38	25.9	739	1	HYPF_RHOCA	Q02987	rhodobacter
170	39	26.5	682	1	RPOC_ORYSA	P12092	oryza sativ	243	38	25.9	779	1	Y100_MYCTU	Q10821	mycobacteri
171	39	26.5	683	1	RPOC_MAIZE	P16024	zea mays (m	244	38	25.9	796	1	MEN_DROME	P52302	drosophila
172	39	26.5	697	1	Y441_HUMAN	O43167	homo sapien	245	38	25.9	805	1	PRIA_BACSU	P94461	bacillus su
173	39	26.5	750	1	HYPF_ECOLI	P30131	escherichia	246	38	25.9	819	1	LON_CHLPN	Q929f4	chlamydia p
174	39	26.5	793	1	SMO_MOUSE	P56726	mus musculus	247	38	25.9	828	1	YFA4_YEAST	P43585	saccharomyc
175	39	26.5	793	1	SMO_RAT	P97698	rattus norv	248	38	25.9	837	1	PHLD_MOUSE	O70362	mus musculus
176	39	26.5	841	1	PHL2_HUMAN	Q15127	homo sapien	249	38	25.9	840	1	PHL1_HUMAN	P80108	homo sapien
177	39	26.5	847	1	STA6_HUMAN	P42226	homo sapien	250	38	25.9	864	1	LCLR_MOUSE	P35951	mus musculus
178	39	26.5	853	1	CIRA_RAT	P15387	rattus norv	251	38	25.9	865	1	TOP1_ECOLI	P06612	escherichia
179	39	26.5	982	1	MSHM_SARGL	O63852	sarcophyton	252	38	25.9	867	1	SYA_AQUAE	O67323	aquifex aeo

253	1	1171	38	25.9	103833	klebsiella	P03833	326	37	25.2	272	1	THIM_BACSU	P39593	bacillus su
254	1	1354	38	25.9	5421	drosophila	P5421	327	37	25.2	284	1	IPYR_PTICPA	O13505	pichia past
255	1	1361	38	25.9	1661	xenopus lae	Q91661	328	37	25.2	287	1	YICC_ECOLI	P23839	escherichia
256	1	1459	38	25.9	1874	caenorhabdi	Q21874	329	37	25.2	289	1	IPYR_BOVIN	P37980	bos taurus
257	1	1675	38	25.9	00610	homo sapien	Q00610	330	37	25.2	296	1	YC32_METJA	Q58629	methanococc
258	1	1675	38	25.9	9951	bos taurus	P49951	331	37	25.2	297	1	ACTR_BOVIN	P34974	bos taurus
259	1	1675	38	25.9	11442	rattus norv	P11442	332	37	25.2	316	1	LKHA_DICDI	P52922	dictyosteli
260	1	1681	38	25.9	34574	caenorhabdi	P34574	333	37	25.2	320	1	MEC3_CAEVU	P34765	caenorhabdi
261	1	1685	38	25.9	29400	homo sapien	P29400	334	37	25.2	321	1	MEC3_CAEEL	P09088	caenorhabdi
262	1	1788	38	25.9	09221	caenorhabdi	Q09221	335	37	25.2	323	1	YBR5_BORPE	Q30446	bordetella
263	1	1852	38	25.9	22316	cyprinus ca	P22316	336	37	25.2	324	1	FASA_RAT	Q63199	rattus norv
264	1	1895	38	25.9	14951	caenorhabdi	P14951	337	37	25.2	327	1	GDB2_WHEAT	P08453	triticult ae
265	1	2064	38	25.9	91621	drosophila	P91621	338	37	25.2	327	1	VMP_CAMVN	Q00966	cauliflower
266	1	2095	38	25.9	37800	toscana vir	P37800	339	37	25.2	329	1	GC2_CAVPO	P01862	cavia porce
267	1	2437	38	25.9	46530	brachydanio	P46530	340	37	25.2	333	1	YZ37_SYNY3	Q55480	synecocyst
268	1	2476	38	25.9	09873	homo sapien	Q09873	341	37	25.2	335	1	CHM1_BOVIN	P17404	bos taurus
269	1	2476	38	25.9	28983	sus scrofa	Q28983	342	37	25.2	345	1	APOH_MOUSE	Q01339	mus musculus
270	1	2813	38	25.9	04275	homo sapien	P04275	343	37	25.2	368	1	LNK_RAT	P50745	rattus norv
271	1	2871	38	25.9	35555	homo sapien	P35555	344	37	25.2	375	1	CD14_HUMAN	P08571	homo sapien
272	1	3033	38	25.9	6660	h genome po	P26660	345	37	25.2	376	1	ADH1_RHOSH	P72324	rhodobacter
273	1	3033	38	25.9	6661	h genome po	P26661	346	37	25.2	376	1	CGD3_ARATH	P42753	arabidopsis
274	1	3951	38	25.9	7920	avian infec	P27920	347	37	25.2	387	1	D4DR_RAT	P30729	rattus norv
275	37.5	48	37.5	25.5	16895	anemonia su	P16895	348	37	25.2	410	1	YG31_YEAST	P53286	saccharomyc
276	37.5	52	37.5	25.5	2253	scythrops n	P52253	349	37	25.2	424	1	IHBA_MOUSE	Q04998	mus musculus
277	37.5	54	37.5	25.5	5608	argusianus	P05608	350	37	25.2	424	1	IHBA_PIG	P03970	sus scrofa
278	37.5	54	37.5	25.5	2239	balearica p	P52239	351	37	25.2	424	1	IHBA_RAT	P18331	rattus norv
279	37.5	54	37.5	25.5	5612	grus carunc	P05612	352	37	25.2	425	1	IHBA_BOVIN	P07995	bos taurus
280	37.5	180	37.5	25.5	5608	homo sapien	P55608	353	37	25.2	425	1	IHBA_SHEEP	P43032	ovis aries
281	37.5	210	37.5	25.5	5185	drosophila	P54185	354	37	25.2	426	1	IHBA_HUMAN	P08476	homo sapien
282	37.5	236	37.5	25.5	7855	oryctolagus	P47855	355	37	25.2	447	1	GNT2_HUMAN	Q10469	h alpha-1,6
283	37.5	238	37.5	25.5	14138	homo sapien	P14138	356	37	25.2	459	1	RSP6_CHLRE	Q01657	chlamydomon
284	37.5	285	37.5	25.5	06990	papio papio	Q06990	357	37	25.2	471	1	NRAM_IAPFW	P18881	influenza a
285	37.5	375	37.5	25.5	99527	homo sapien	Q099527	358	37	25.2	475	1	SYL_BOVIN	P17248	bos taurus
286	37.5	375	37.5	25.5	00600	xenopus lae	P70060	359	37	25.2	483	1	MYPE_MOUSE	P70402	mus musculus
287	37.5	384	37.5	25.5	4838	rattus norv	O54838	360	37	25.2	485	1	CAT1_NICPL	P49315	nicotiana p
288	37.5	424	37.5	25.5	3611	mus musculus	P23611	361	37	25.2	492	1	SYVM_YEAST	P48527	saccharomyc
289	37.5	484	37.5	25.5	5835	mycobacteri	O05835	362	37	25.2	492	1	YEOL_YEAST	P40034	saccharomyc
290	37.5	493	37.5	25.5	15048	homo sapien	Q15048	363	37	25.2	497	1	TRPE_ACICA	P23315	acinetobact
291	37.5	498	37.5	25.5	2905	aspergillus	Q02905	364	37	25.2	501	1	DLDH_PEA	P31023	plisum sativ
292	37.5	499	37.5	25.5	10529	aspergillus	P10529	365	37	25.2	509	1	SYK_ACICA	Q43990	acinetobact
293	37.5	499	37.5	25.5	2906	aspergillus	Q02906	366	37	25.2	512	1	DLDH_SCHPO	O00087	schizosacch
294	37.5	499	37.5	25.5	30292	aspergillus	P30292	367	37	25.2	548	1	ERE_HUMAN	P50548	homo sapien
295	37.5	562	37.5	25.5	00750	homo sapien	P00750	368	37	25.2	551	1	ERE_MOUSE	P70459	mus musculus
296	37.5	568	37.5	25.5	48998	rattus norv	P48998	369	37	25.2	584	1	SYR_SYNY3	Q55486	synecocyst
297	37.5	590	37.5	25.5	5479	bacterioph	P25479	370	37	25.2	608	1	ALBU_RABIT	P49065	oryctolagus
298	37.5	640	37.5	25.5	6445	rauscher m	P06445	371	37	25.2	609	1	PRXC_CURIN	P49053	curvularia
299	37.5	643	37.5	25.5	25481	human papil	P25481	372	37	25.2	623	1	NRC2_NEUCR	O42626	neurospora
300	37.5	776	37.5	25.5	40596	azotobacter	P40596	373	37	25.2	662	1	Y4NK_RHISN	P55583	rhizobium s
301	37.5	849	37.5	25.5	170175	mus musculus	P70175	374	37	25.2	691	1	TALA_POVBA	P14999	polyomaviru
302	37.5	849	37.5	25.5	62936	rattus norv	Q62936	375	37	25.2	695	1	IE63_MCMVS	P03071	polyomaviru
303	37.5	992	37.5	25.5	13567	micrococcus	P13567	376	37	25.2	841	1	NEXR_RAT	Q69154	murine cyto
304	37.5	1125	37.5	25.5	3992	homo sapien	P53992	377	37	25.2	863	1	MIS5_SCHPO	P15205	rattus norv
305	37.5	1191	37.5	25.5	34203	african swi	P34203	378	37	25.2	868	1	SYA_AQUPY	P49731	schizosacch
306	37.5	1192	37.5	25.5	00942	herpes simp	P00942	379	37	25.2	871	1	SYA_AQUPY	Q9xdm3	aquifex pyr
307	37.5	1196	37.5	25.5	89452	herpes simp	P89452	380	37	25.2	894	1	SYLM_SACDO	P13503	saccharomyc
308	37.5	1197	37.5	25.5	36384	herpes simp	Q63684	381	37	25.2	896	1	Y249_HUMAN	Q92539	homo sapien
309	37.5	1210	37.5	25.5	10279	mus musculus	P010279	382	37	25.2	1046	1	POL_SIVAG	P27980	simian immu
310	37.5	1298	37.5	25.5	39523	haemophilus	Q57523	383	37	25.2	1046	1	RPOC_WBIHE	P96177	weissella h
311	37	124	37	25.2	09304	caenorhabdi	Q09304	388	37	25.2	1240	1	POL_SIVAT	P27973	simian immu
312	37	138	37	25.2	1223	bos taurus	P01223	390	37	25.2	1376	1	POL_SIVAT	P27973	simian immu
313	37	138	37	25.2	4828	canis famil	P54828	391	37	25.2	1394	1	RPOC_WEIPA	P05895	simian immu
314	37	180	37	25.2	3618	oryctolagus	P33618	392	37	25.2	1437	1	POL_EEEV	P08768	eastern equ
315	37	207	37	25.2	16455	homo sapien	P16455	393	37	25.2	1443	1	POL_EEEV3	P27284	eastern equ
316	37	212	37	25.2	5207	haemophilus	P45207	394	37	25.2	1468	1	VGL2_CVM4	P22432	murine coro
317	37	252	37	25.2	20851	homo sapien	P20851	395	37	25.2	1616	1	VGL2_CVM4	P22432	murine coro
318	37	252	37	25.2	1223	bos taurus	P01223	399	37	25.2	1376	1	VGL2_CVM4	P22432	murine coro
319	37	252	37	25.2	4828	canis famil	P54828	399	37	25.2	1394	1	VGL2_CVM4	P22432	murine coro
320	37	252	37	25.2	3618	oryctolagus	P33618	399	37	25.2	1437	1	VGL2_CVM4	P22432	murine coro
321	37	252	37	25.2	16455	homo sapien	P16455	399	37	25.2	1443	1	VGL2_CVM4	P22432	murine coro
322	37	252	37	25.2	5207	haemophilus	P45207	399	37	25.2	1468	1	VGL2_CVM4	P22432	murine coro
323	37	252	37	25.2	20851	homo sapien	P20851	399	37	25.2	1616	1	VGL2_CVM4	P22432	murine coro
324	37	252	37	25.2	1223	bos taurus	P01223	399	37	25.2	1376	1	VGL2_CVM4	P22432	murine coro
325	37	252	37	25.2	4828	canis famil	P54828	399	37	25.2	1394	1	VGL2_CVM4	P22432	murine coro

399	1	MAPB_HUMAN	2468	25.2	37	P46821	472	36	24.5	290	1	HYPB_ECOLI	P24190	escherichia
400	1	NTC1_RAT	2531	25.2	37	Q07008	473	36	24.5	302	1	YE86_MYCTU	P71766	mycobacteri
401	1	FBN1_BOVIN	2871	25.2	37	P98133	474	36	24.5	307	1	GDA9_WHEAT	P18573	tritiscum ae
402	1	CA1C_MOUSE	3067	25.2	37	Q60847	475	36	24.5	309	1	41BL_MOUSE	P42274	mus musculus
403	1	LMAG2_HUMAN	3110	25.2	37	P24043	476	36	24.5	309	1	MAT1_HUMAN	P51948	homo sapien
404	1	TEGU_HSV11	3164	25.2	37	P10220	477	36	24.5	309	1	MAT1_MOUSE	P51948	mus musculus
405	1	IOVO_CALSCP	56	24.8	36.5	P05589	478	36	24.5	309	1	MAT1_XENLA	P51951	xenopus lae
406	1	IOVO_CALSCP	56	24.8	36.5	P05588	479	36	24.5	314	1	MAT1_MOUSE	Q63844	mus musculus
407	1	MT1_COLLI	63	24.8	36.5	P15786	480	36	24.5	320	1	YE86_MYCAV	Q63844	mus musculus
408	1	ARSR_STAAU	104	24.8	36.5	P30338	481	36	24.5	323	1	KRAF_MSV36	O07402	mycobacteri
409	1	ARSR_STAAU	104	24.8	36.5	Q01256	482	36	24.5	333	1	KRAF_MSV36	P00532	murine sarc
410	1	IF5_TRICA	162	24.8	36.5	Q26891	483	36	24.5	340	1	GLB_PSEDC	P26914	pseudoterra
411	1	EMBI1_CAVPO	233	24.8	36.5	P22032	484	36	24.5	356	1	AC13_HUMAN	Q05127	ebola virus
412	1	KAF2_SORBI	267	24.8	36.5	P14691	485	36	24.5	356	1	GLTA_WHEAT	P40938	homo sapien
413	1	IBP2_RAT	304	24.8	36.5	P12843	486	36	24.5	363	1	TOB_MOUSE	P10385	tritiscum ae
414	1	IBP2_SHEEP	317	24.8	36.5	Q29400	487	36	24.5	363	1	ROX1_YEAST	Q61471	mus musculus
415	1	OM52_HAEIN	333	24.8	36.5	P38368	488	36	24.5	369	1	HEM3_PEA	P25042	saccharomyc
416	1	IF5_RAT	429	24.8	36.5	Q07205	489	36	24.5	380	1	MK03_RAT	Q43082	pisum sativ
417	1	IF5_HUMAN	431	24.8	36.5	P55010	490	36	24.5	386	1	VAT2_GOSHI	P21708	r mitogen-a
418	1	YN95_YEAST	436	24.8	36.5	P53752	491	36	24.5	389	1	RIR2_HUMAN	Q43433	gossypium h
419	1	AP2_HUMAN	437	24.8	36.5	P05549	492	36	24.5	391	1	PEXG_YARLI	P31350	homo sapien
420	1	AP2_MOUSE	437	24.8	36.5	P34056	493	36	24.5	404	1	CAG5_CHICK	P78980	yarrowia li
421	1	EX1_HAEIN	473	24.8	36.5	P45188	494	36	24.5	418	1	PEDE_HUMAN	Q92184	gallus gall
422	1	NUSA_BOYBU	482	24.8	36.5	Q51740	495	36	24.5	426	1	CNAL_CANAL	P32782	candida alb
423	1	GATI1_YEAST	510	24.8	36.5	P43574	496	36	24.5	430	1	FTSZ_SYN3	P73456	synchocyst
424	1	INR2_SHEEP	536	24.8	36.5	Q95207	497	36	24.5	431	1	PKNA_MYCTU	P71585	mycobacteri
425	1	YKCS_HUMAN	541	24.8	36.5	P41996	498	36	24.5	432	1	YK27_CABEL	O16686	caenorhabdi
426	1	CSF1_HUMAN	554	24.8	36.5	P09603	499	36	24.5	446	1	SH7_CAVPO	P50407	cavia porce
427	1	CNAL_DROME	584	24.8	36.5	P12252	500	36	24.5	459	1	BGLS_AGRSP	P12614	agrobacteri
428	1	VE1_HPV27	643	24.8	36.5	P36723	501	36	24.5	461	1	CGRR_HUMAN	Q16602	homo sapien
429	1	VE1_HPV57	643	24.8	36.5	P22153	502	36	24.5	464	1	CGRR_RAT	Q63118	rattus norv
430	1	TF1B_MOUSE	834	24.8	36.5	Q62318	503	36	24.5	469	1	VL2_HPV35	P27234	human papil
431	1	MCN2_DROME	887	24.8	36.5	P49735	504	36	24.5	471	1	SYW_HUMAN	P23381	homo sapien
432	1	GCSP_PEA	1037	24.8	36.5	P26969	505	36	24.5	475	1	SYW_RABIT	P23612	oryctolagus
433	1	SYI_TETH	1081	24.8	36.5	P36422	506	36	24.5	479	1	ATIN_HSV1F	P04486	herpes simp
434	1	CC4H_HUMAN	1918	24.8	36.5	P50851	507	36	24.5	481	1	SYW_MOUSE	P32921	mus musculus
435	1	TOR1_YEAST	2470	24.8	36.5	P35169	508	36	24.5	490	1	ATIN_HSV11	P06492	herpes simp
436	1	TXOA_HADVE	37	24.5	36	P56207	509	36	24.5	495	1	MLP2_DROME	Q24400	drosophila
437	1	TXOC_HADVE	37	24.5	36	P81596	510	36	24.5	496	1	G6PD_THEMA	Q9X0N9	thermotoga
438	1	RUBR_CLOTS	52	24.5	36	P19500	511	36	24.5	515	1	CP51_PENIT	Q26664	penicillium
439	1	RUBR_ACICA	54	24.5	36	P42453	512	36	24.5	542	1	ZYX_CHICK	Q40584	gallus gall
440	1	RL32_TETH	59	24.5	36	P80339	513	36	24.5	570	1	HAO_NITEU	O50925	nitrosomona
441	1	VE7_HPV1A	93	24.5	36	P06465	514	36	24.5	579	1	YD49_MYCTU	Q11019	mycobacteri
442	1	Y04N_BPT4	104	24.5	36	P39233	515	36	24.5	582	1	YANB_SCHPO	Q10076	schizosacch
443	1	VMEN_PMV	111	24.5	36	P20953	516	36	24.5	587	1	TPYA_SYN3	P72749	synchocyst
444	1	FER_PSEAE	114	24.5	36	Q51383	517	36	24.5	598	1	NR42_RAT	Q07917	rattus norv
445	1	IPGE_SHIFL	120	24.5	36	Q07567	518	36	24.5	603	1	HIP1_YEAST	P06775	saccharomyc
446	1	GLHA_FUNHE	125	24.5	36	P47744	519	36	24.5	605	1	UGST_MAIZE	P04713	zea mays (m
447	1	KRCL_CHICK	127	24.5	36	P25692	520	36	24.5	617	1	UGST_SORBI	Q43134	sorghum bic
448	1	TSHE_LAMGL	138	24.5	36	P79357	521	36	24.5	622	1	PPFA_RICCO	Q41140	ricinus com
449	1	TSHE_PIG	138	24.5	36	P01224	522	36	24.5	622	1	MAK_MOUSE	Q04859	mus musculus
450	1	FUR1_BACSU	145	24.5	36	P54479	523	36	24.5	629	1	SYM_THEMA	O33925	thermotoga
451	1	SODC_CANAL	153	24.5	36	Q59924	524	36	24.5	648	1	KRAF_HUMAN	P04049	homo sapien
452	1	GP27_BPSP1	155	24.5	36	P06228	525	36	24.5	648	1	KRAF_RAT	P11345	rattus norv
453	1	NCS2_CABEL	189	24.5	36	P36609	526	36	24.5	665	1	ENV_MLYMO	P03385	moloney mur
454	1	Y335_MYCPN	193	24.5	36	P75303	527	36	24.5	667	1	KJF7_YEAST	P47042	saccharomyc
455	1	RK24_PEA	194	24.5	36	P11893	528	36	24.5	667	1	SIM2_HUMAN	Q14190	homo sapien
456	1	MGMT_CRIGR	209	24.5	36	P26186	529	36	24.5	667	1	TS11_GIALA	Q03185	giardia lam
457	1	Y347_MYCPN	210	24.5	36	P75256	530	36	24.5	727	1	VP4_RDV	P22474	rice dwarf
458	1	CTLA_HUMAN	223	24.5	36	P16410	531	36	24.5	758	1	YIDM_CABEL	Q94225	caenorhabdi
459	1	CTLA_MOUSE	223	24.5	36	P09793	532	36	24.5	771	1	ITB2_MOUSE	P18335	mus musculus
460	1	CTLA_RABIT	223	24.5	36	P42072	533	36	24.5	785	1	VD05_VACCV	P04305	vaccinia vi
461	1	VGLL_HSV2H	224	24.5	36	P28278	534	36	24.5	802	1	CIKB_RAT	Q63099	rattus norv
462	1	YKPI_KLULA	224	24.5	36	P05467	535	36	24.5	806	1	CIKB_HUMAN	Q92953	homo sapien
463	1	CDSN_PIG	225	24.5	36	O19084	536	36	24.5	816	1	NPA2_MOUSE	P97460	mus musculus
464	1	KLK_PIG	232	24.5	36	P00752	537	36	24.5	819	1	LON_CHLZR	O84348	chlamydia t
465	1	DNAB_MYCLE	237	24.5	36	P46394	538	36	24.5	824	1	MS2_HUMAN	P78325	homo sapien
466	1	LLDR_ECOLI	238	24.5	36	P33233	539	36	24.5	836	1	GYRA_MYCGE	P80109	bos taurus
467	1	ALDC_KLETE	259	24.5	36	Q04518	540	36	24.5	839	1	PHLD_BOVIN	P42500	mycoplasma
468	1	RS2_BOYBU	260	24.5	36	Q51149	541	36	24.5	841	1	TTK_HUMAN	P33981	homo sapien
469	1	YIGW_ECOLI	264	24.5	36	P27859	542	36	24.5	878	1	SV4_BACSU	O34526	bacillus su
470	1	VG11_HSV11	281	24.5	36	Q00164	543	36	24.5	943	1	CENC_HUMAN	Q03188	homo sapien
471	1	HMD4_XENLA	285	24.5	36	P53775	544	36	24.5	963	1	YO36_CAEEL	Q09457	caenorhabdi

545	36	24.5	974	1	YMB4_CAEEL	Q03601 caenorhabdi	618	35.5	24.1	4092	1	DYHC_YEAST	P36022 saccharomyc
546	36	24.5	976	1	EPA2_HUMAN	P29317 homo sapien	619	35	23.8	37	1	TXOB_HADVE	P81595 hadronyche
547	36	24.5	1020	1	GCSP_HUMAN	P23378 homo sapien	620	35	23.8	37	1	TXOB_HADVE	P81597 hadronyche
548	36	24.5	1035	1	ITA9_HUMAN	Q13797 homo sapien	621	35	23.8	37	1	TXOE_HADVE	P81598 hadronyche
549	36	24.5	1059	1	SPS_VICFA	Q43876 vicia faba	622	35	23.8	52	1	RUBR_DESGI	P00270 desulfovibr
550	36	24.5	1061	1	RNE_ECOLI	P21513 escherichia	623	35	23.8	105	1	HOR7_HORVU	P06472 hordeum vul
551	36	24.5	1115	1	PHYE_PHANI	P55004 pharbitis n	624	35	23.8	115	1	YPOL_IBDVP	P25222 avian infec
552	36	24.5	1163	1	RPOD_PEA	P12227 pisum sativ	625	35	23.8	138	1	TSHB_HORSE	Q28376 equus caball
553	36	24.5	1173	1	TSP1_XENLA	P35448 xenopus lae	626	35	23.8	141	1	LSHB_PIG	P01232 sus scrofa
554	36	24.5	1202	1	NS03_HUMAN	P29474 homo sapien	627	35	23.8	145	1	YPOL_IBDVS	P25221 avian infec
555	36	24.5	1207	1	EGF_HUMAN	P09133 homo sapien	628	35	23.8	145	1	YPOL_IBDVC	P15481 avian infec
556	36	24.5	1221	1	FB12_MOUSE	P37889 mus musculu	629	35	23.8	145	1	YPOL_IBDVS	P22440 avian infec
557	36	24.5	1281	1	YLB5_CAEEL	P46580 caenorhabdi	630	35	23.8	151	1	SC1D_EMENI	P19815 emericella
558	36	24.5	1415	1	LYS2_SCHPO	P40976 schizosacch	631	35	23.8	154	1	KRSC_CHICK	P04459 gallus gall
559	36	24.5	1435	1	DPO3_MYCPU	P47729 mycoplasma	632	35	23.8	181	1	RG79_HSV1	Q00148 ictaluriid h
560	36	24.5	1451	1	DPO3_MYCPE	P47277 mycoplasma	633	35	23.8	186	1	GDA8_WHEAT	P04728 triticum ae
561	36	24.5	1480	1	PAN1_YEAST	P32521 saccharomyc	634	35	23.8	186	1	GL19_ORYSA	P29835 oryza sativ
562	36	24.5	1534	1	MTDM_ARATH	P34881 arabidopsis	635	35	23.8	190	1	PGHD_HUMAN	P41222 homo sapien
563	36	24.5	1597	1	M3K4_MOUSE	O08648 mus musculu	636	35	23.8	202	1	LEXA_ERWCA	O04596 erwinia car
564	36	24.5	1607	1	LMG1_MOUSE	P02468 mus musculu	637	35	23.8	210	1	IF2G_SPIVO	Q36041 spironucleu
565	36	24.5	1609	1	LMG1_HUMAN	P11047 homo sapien	638	35	23.8	219	1	BTE2_HUMAN	Q13887 homo sapien
566	36	24.5	1613	1	VIT2_CAEEL	P05690 caenorhabdi	639	35	23.8	228	1	LITF_HUMAN	Q99732 homo sapien
567	36	24.5	1720	1	F7SH_CHLYU	P56369 chlorella v	640	35	23.8	229	1	VG14_BPPH8	P14814 bacterioph
568	36	24.5	1799	1	YO25_CAEEL	P34675 caenorhabdi	641	35	23.8	236	1	RVVA_DABRU	P18964 daboia russ
569	36	24.5	1871	1	PLX4_HUMAN	P51805 homo sapien	642	35	23.8	236	1	RVVG_DABRU	P18965 daboia russ
570	36	24.5	2318	1	NTC3_MOUSE	O61982 mus musculu	643	35	23.8	237	1	NRL_HUMAN	P54845 homo sapien
571	36	24.5	2524	1	NOTC_XENLA	P21783 xenopus lae	644	35	23.8	237	1	NRL_MOUSE	P54846 mus musculu
572	36	24.5	2616	1	NDL_DROME	P98159 drosophila	645	35	23.8	249	1	NODH_RHTR	P52994 thizobium t
573	36	24.5	3418	1	BRC2_HUMAN	P51587 homo sapien	646	35	23.8	250	1	FATX_BOTAT	P04971 bothriops at
574	36	24.5	3866	1	HRX_MOUSE	P55200 mus musculu	647	35	23.8	260	1	FLVB_TRIFL	P05620 trimeresuru
575	36	24.5	6486	1	TYCC_BACBR	O30409 b tyrocidin	648	35	23.8	262	1	GDA1_WHEAT	P04721 triticum ae
576	35.5	24.1	51	1	IOVO_RHYFU	P52251 rhynchotus	649	35	23.8	266	1	ELA_ADES7	P06499 simian aden
577	35.5	24.1	54	1	IOVO_PAVMU	P52263 pavo muticu	650	35	23.8	266	1	STR_KLEPN	P13082 kiebsiella
578	35.5	24.1	56	1	IOVO_AFRCO	P52258 afroapo co	651	35	23.8	267	1	CYSE_HAEIN	P43886 haemophilus
579	35.5	24.1	56	1	IOVO_PAVCR	P05609 pavo crista	652	35	23.8	272	1	HAT5_ARATH	Q02283 arabidopsis
580	35.5	24.1	171	1	NUPM_BOVIN	P42029 bos taurus	653	35	23.8	282	1	NADA_APLKU	Q27312 aplysia kur
581	35.5	24.1	189	1	INAD_BOVIN	P07348 bos taurus	654	35	23.8	286	1	GDA0_WHEAT	P02863 triticum ae
582	35.5	24.1	189	1	INAD_BOVIN	P05010 bos taurus	655	35	23.8	289	1	FRAH_ANASP	P46017 anabaena sp
583	35.5	24.1	210	1	IOVO_CHICK	P10005 gallus gall	656	35	23.8	289	1	HOG3_HORVU	P80198 hordeum vul
584	35.5	24.1	250	1	HXB9_MOUSE	P20615 mus musculu	657	35	23.8	289	1	YEC4_EBV	P03235 epstein-bar
585	35.5	24.1	257	1	CLCA_RHOOP	O67987 rhodococcus	658	35	23.8	291	1	GDA2_WHEAT	P04722 triticum ae
586	35.5	24.1	285	1	YKG2_CAEEL	P46552 caenorhabdi	659	35	23.8	291	1	US02_HSV11	P06485 herpes simp
587	35.5	24.1	311	1	MDH_HAEIN	P44427 haemophilus	660	35	23.8	291	1	US02_HSV2H	P13292 herpes simp
588	35.5	24.1	325	1	Y272_SYNY3	P73893 synechocyst	661	35	23.8	297	1	GDA4_WHEAT	P04724 triticum ae
589	35.5	24.1	375	1	FMOD_BOVIN	P13605 bos taurus	662	35	23.8	307	1	GLTB_WHEAT	P10386 triticum ae
590	35.5	24.1	376	1	FMOD_MOUSE	P50608 mus musculu	663	35	23.8	326	1	UL31_HSVEB	P28951 equine herp
591	35.5	24.1	376	1	FMOD_RAT	P50609 rattus norv	664	35	23.8	326	1	YJHS_ECOLI	P29370 escherichia
592	35.5	24.1	407	1	YO14_CAEEL	P43570 caenorhabdi	665	35	23.8	329	1	PIN2_CAEEL	Q19157 caenorhabdi
593	35.5	24.1	469	1	NRAM_TAHCO	P08327 influenza a	666	35	23.8	345	1	APDH_HUMAN	P02749 homo sapien
594	35.5	24.1	527	1	RAG2_HUMAN	P55895 homo sapien	667	35	23.8	353	1	IL8B_MACMU	Q28519 macaca mula
595	35.5	24.1	527	1	RAG2_MOUSE	P21784 mus musculu	668	35	23.8	358	1	CCDA_CAEEL	P35800 caenorhabdi
596	35.5	24.1	527	1	RAG2_RABIT	P34089 oryctolagus	669	35	23.8	358	1	WNT8_XENLA	P28026 xenopus lae
597	35.5	24.1	556	1	MP1P_EMENI	P30303 emericella	670	35	23.8	370	1	CPR3_CAEEL	P43507 caenorhabdi
598	35.5	24.1	578	1	NRD2_RAT	Q63504 rattus norv	671	35	23.8	378	1	ADH3_HAEIN	P44557 haemophilus
599	35.5	24.1	649	1	SYM_BACST	P23920 bacillus st	672	35	23.8	380	1	KMIL_AVTMH	P00531 avian retro
600	35.5	24.1	673	1	IP3L_RAT	P42335 rattus norv	673	35	23.8	384	1	YEHL_ECOLI	P33348 escherichia
601	35.5	24.1	674	1	RNE_SYNY3	P72656 synechocyst	674	35	23.8	388	1	CHSD_PHANI	Q22045 pharbitis n
602	35.5	24.1	698	1	TRPE_RAT	P12346 rattus norv	675	35	23.8	390	1	NEUC_ECOLI	Q24700 escherichia
603	35.5	24.1	730	1	BP1_HUMAN	P13497 homo sapien	676	35	23.8	391	1	GA6A_XENLA	Q91678 xenopus lae
604	35.5	24.1	757	1	ASPH_HUMAN	Q12797 homo sapien	677	35	23.8	391	1	GA6B_XENLA	P70005 xenopus lae
605	35.5	24.1	779	1	M130_STRPU	P08472 strongyloce	678	35	23.8	393	1	YCFU_HAEIN	P44252 haemophilus
606	35.5	24.1	917	1	SYM_CAEEL	Q20970 caenorhabdi	679	35	23.8	394	1	ARAJ_ECOLI	P23910 escherichia
607	35.5	24.1	925	1	PIP1_YEAST	P40020 saccharomyc	680	35	23.8	399	1	IRTF_MOUSE	Q61179 mus musculu
608	35.5	24.1	925	1	UVRA_ZYMMO	O31151 zymomonas m	681	35	23.8	400	1	PRTZ_HUMAN	P22891 homo sapien
609	35.5	24.1	991	1	BP1_MOUSE	P98063 mus musculu	682	35	23.8	402	1	CENC_SHEEP	P49453 ovis aries
610	35.5	24.1	1194	1	MGR1_HUMAN	Q13255 homo sapien	683	35	23.8	409	1	KPRS_SCHPO	P41831 schizosacch
611	35.5	24.1	1242	1	IRS1_HUMAN	P35568 homo sapien	684	35	23.8	415	1	CXA3_RAT	P29414 rattus norv
612	35.5	24.1	1427	1	REST_HUMAN	P30622 homo sapien	685	35	23.8	428	1	VMSA_HPBGS	P03144 ground squi
613	35.5	24.1	1863	1	BRCL_HUMAN	P38398 homo sapien	686	35	23.8	428	1	WN8B_XENLA	P31291 xenopus lae
614	35.5	24.1	1878	1	BRCL_CANFA	O95153 canis famil	687	35	23.8	430	1	NOR2_RAT	Q63516 rattus norv
615	35.5	24.1	2481	1	UN52_CAEEL	Q06561 caenorhabdi	688	35	23.8	434	1	CBPE_BOVIN	P04836 bos taurus
616	35.5	24.1	2907	1	FBN2_MOUSE	Q01555 mus musculu	689	35	23.8	441	1	YLIC_ECOLI	P75802 escherichia
617	35.5	24.1	2911	1	FBN2_HUMAN	P35556 homo sapien	690	35	23.8	453	1	BIOA_AQUAE	O66557 aquifex aeo

691	35	23.8	455	1	RADA_TREPA	083985	treponema p	764	35	23.8	1073	1	RAG1_ONCMY	Q91187	oncorhynchu
692	35	23.8	461	1	TSPI_BOVIN	Q28178	bos taurus	765	35	23.8	1085	1	YAF4_SCHPO	Q09863	schizosacch
693	35	23.8	476	1	CBPH_HUMAN	P16870	homo sapien	766	35	23.8	1093	1	P14K_DICDI	P54677	dictyosteli
694	35	23.8	476	1	CBPH_MOUSE	Q00493	mus musculus	767	35	23.8	1100	1	RAD1_YEAST	P06777	saccharomyc
695	35	23.8	476	1	CBPH_RAT	P15087	rattus norv	768	35	23.8	1128	1	PHYA_ORISA	P10931	oryza sativ
696	35	23.8	486	1	CDSN_HUMAN	Q15517	homo sapien	769	35	23.8	1150	1	2ACA_HUMAN	Q06190	homo sapien
697	35	23.8	499	1	GUNI_BACSU	P07983	bacillus su	770	35	23.8	1169	1	YK82_YEAST	P36170	saccharomyc
698	35	23.8	499	1	YHM8_YEAST	P38860	saccharomyc	771	35	23.8	1170	1	TSPI_HUMAN	P07996	homo sapien
699	35	23.8	500	1	YDGR_ECOLI	P77304	escherichia	772	35	23.8	1170	1	TSPI_MOUSE	P35441	mus musculus
700	35	23.8	506	1	30LD_NOCOP	Q04616	nocardia op	773	35	23.8	1175	1	PNF_RAT	Q62728	rattus norv
701	35	23.8	519	1	TYR2_HUMAN	P40126	homo sapien	774	35	23.8	1176	1	VPS8_YEAST	P39702	saccharomyc
702	35	23.8	527	1	CATA_RAT	P04762	rattus norv	775	35	23.8	1178	1	TSPI_CHICK	P35440	gallus gall
703	35	23.8	528	1	CP51_CANAL	P10613	candida alb	776	35	23.8	1204	1	YAEF_SCHPO	Q09854	schizosacch
704	35	23.8	528	1	CP51_CANTR	P14263	candida tro	777	35	23.8	1295	1	PURL_ECOLI	P15254	escherichia
705	35	23.8	528	1	YNH7_YEAST	P53940	saccharomyc	778	35	23.8	1400	1	RON_HUMAN	Q04912	homo sapien
706	35	23.8	530	1	INR2_BOVIN	Q95141	bos taurus	779	35	23.8	1487	1	MDS3_YEAST	P53094	saccharomyc
707	35	23.8	531	1	PKYK_EIMTE	Q44006	elmeria ten	780	35	23.8	1500	1	SSP5_STRGN	P16952	streptococc
708	35	23.8	537	1	GAG_BAEYM	P03341	baboon endo	781	35	23.8	1574	1	RPOC_AQUAE	Q67763	aquifex aeo
709	35	23.8	537	1	OGP_BOVIN	Q28042	bos taurus	782	35	23.8	1687	1	2142_HUMAN	P52746	homo sapien
710	35	23.8	539	1	OGP_SHEEP	Q28542	ovis aries	783	35	23.8	1691	1	POLN_HEVME	Q03495	hepatitis e
711	35	23.8	554	1	SYNC_YEAST	P38707	saccharomyc	784	35	23.8	1722	1	RBB2_HUMAN	P29375	homo sapien
712	35	23.8	558	1	OL6G_BACCE	P13332	bacillus ce	785	35	23.8	2127	1	RRPL_RABVP	P16289	rabies viru
713	35	23.8	559	1	POUL_DUGJA	P13370	dugesia jap	786	35	23.8	2142	1	RRPL_RABVP	P11213	rabies viru
714	35	23.8	563	1	IDS_MOUSE	Q08890	mus musculus	787	35	23.8	2747	1	FAF_DROME	P34926	rattus norv
715	35	23.8	564	1	GAF2_SCHPO	Q10134	schizosacch	788	35	23.8	2774	1	MAPA_RAT	P53824	drosophila
716	35	23.8	569	1	YAR3_SCHPO	Q10136	schizosacch	789	35	23.8	2805	1	MAPA_HUMAN	P78559	homo sapien
717	35	23.8	575	1	YAJA_BACSU	P45861	bacillus su	790	35	23.8	2813	1	VNF_CANFA	Q28295	canis famill
718	35	23.8	589	1	2132_HUMAN	P52740	homo sapien	791	35	23.8	3106	1	LMA2_MOUSE	Q60675	mus musculus
719	35	23.8	598	1	NR43_HUMAN	P43354	homo sapien	792	35	23.8	3133	1	HMCT_BOMMO	P98092	bombxy mori
720	35	23.8	599	1	HMW3_MYCGE	Q57081	mycoplasma	793	35	23.8	3255	1	POLG_LMVO	P31999	1 genome po
721	35	23.8	603	1	UGST_HORVU	P09842	hordeum vul	794	35	23.8	3255	1	POLG_LMVE	P89876	1 genome po
722	35	23.8	612	1	LEM2_MOUSE	Q00690	mus musculus	795	35	23.8	5147	1	FAT_DROME	P33450	drosophila
723	35	23.8	612	1	MCR_XENLA	Q91573	xenopus lae	796	35	23.8	6359	1	BACC_BACLI	O68008	b bacitraci
724	35	23.8	619	1	KNH2_BOVIN	P01045	bos taurus	797	34.5	23.5	54	1	IOVO_CARRE	P03616	carpococcys
725	35	23.8	619	1	VHPI_CAEEL	Q10038	caenorhabdi	798	34.5	23.5	54	1	IOVO_DENEY	P05569	dendrocryna
726	35	23.8	621	1	KNH1_BOVIN	P01044	bos taurus	799	34.5	23.5	62	1	RS27_METJA	P54028	methanococc
727	35	23.8	626	1	NR43_HUMAN	Q92570	homo sapien	800	34.5	23.5	129	1	RPOA_SCEOB	P19587	scenedesmus
728	35	23.8	628	1	NR43_RAT	P51179	rattus norv	801	34.5	23.5	161	1	ASFL_HELAN	P23257	helianthus
729	35	23.8	639	1	YAJ9_SCHPO	Q09909	schizosacch	802	34.5	23.5	164	1	COX4_SCHPO	P78010	schizosacch
730	35	23.8	647	1	KMIL_CHICK	P05625	gallus gall	803	34.5	23.5	179	1	IECI_LERYVA	P34952	erythrina v
731	35	23.8	649	1	PRTS_MACMO	Q28520	macaca mula	804	34.5	23.5	187	1	EPF_HAEIN	P43771	haemophilus
732	35	23.8	661	1	ENV_MLVCB	P08360	cas-br-e mu	805	34.5	23.5	221	1	ASP_THECC	P32765	theobroma c
733	35	23.8	665	1	PD12_RAT	P20717	rattus norv	806	34.5	23.5	250	1	CD27_MOUSE	P41272	mus musculus
734	35	23.8	671	1	VNCS_PAVHB	P07298	human parvo	807	34.5	23.5	303	1	DMA_TREPA	Q33844	treponema p
735	35	23.8	672	1	SL52_HUMAN	P31639	homo sapien	808	34.5	23.5	306	1	YK84_YEAST	P36171	saccharomyc
736	35	23.8	673	1	PD12_MOUSE	Q08642	mus musculus	809	34.5	23.5	315	1	GBF1_ARATH	P42774	arabidopsis
737	35	23.8	700	1	MEPB_HUMAN	P16820	homo sapien	810	34.5	23.5	331	1	APAZ_KLULA	P49348	kluveromyc
738	35	23.8	717	1	CYG5_HUMAN	Q02108	homo sapien	811	34.5	23.5	333	1	SPT2_YEAST	P06843	saccharomyc
739	35	23.8	718	1	EXT2_HUMAN	Q93063	homo sapien	812	34.5	23.5	338	1	SED1_YEAST	Q01589	saccharomyc
740	35	23.8	739	1	EXT2_MOUSE	P70428	mus musculus	813	34.5	23.5	340	1	GBX2_CHICK	O42230	gallus gall
741	35	23.8	722	1	GUNF_CLOCE	P37698	clostridium	814	34.5	23.5	347	1	MDHM_EUCGU	P45487	eucalyptus
742	35	23.8	767	1	PUBL_SCHPO	Q92462	schizosacch	815	34.5	23.5	359	1	OM53_HAEIN	P45996	haemophilus
743	35	23.8	769	1	ITB2_HUMAN	P05107	homo sapien	816	34.5	23.5	370	1	SEA_AVJET	P23049	avian eryth
744	35	23.8	785	1	VD05_VAVV	P33069	variola vir	817	34.5	23.5	402	1	VGLD_PVTRI	P07645	pseudorabie
745	35	23.8	787	1	ITB3_MOUSE	Q54890	mus musculus	818	34.5	23.5	404	1	SGJ1_YEAST	P25303	saccharomyc
746	35	23.8	806	1	HYF2_ALCEU	Q07451	alcaligenes	819	34.5	23.5	408	1	ENV_RSFFV	P25389	rauscher sp
747	35	23.8	839	1	YNEI_YEAST	P33959	saccharomyc	820	34.5	23.5	409	1	TRAI_MOUSE	P39428	mus musculus
748	35	23.8	844	1	BRO1_YEAST	P48582	saccharomyc	821	34.5	23.5	441	1	YK08_MYCTU	Q10849	mycobacteri
749	35	23.8	847	1	CD22_HUMAN	P20273	homo sapien	822	34.5	23.5	444	1	YA7C_SCHPO	Q09765	schizosacch
750	35	23.8	858	1	UBP5_HUMAN	P45974	homo sapien	823	34.5	23.5	454	1	CG21_SOYBN	P25011	glycine max
751	35	23.8	858	1	UBP5_MOUSE	P56399	mus musculus	824	34.5	23.5	465	1	YQ18_CAEEL	Q09280	caenorhabdi
752	35	23.8	885	1	ITH3_HUMAN	Q60633	homo sapien	825	34.5	23.5	473	1	CG21_ANTWA	P34800	antirrhinum
753	35	23.8	886	1	ITH3_MOUSE	P97280	mesocricetu	826	34.5	23.5	480	1	KRAC_HUMAN	P31750	homo sapien
754	35	23.8	887	1	ITH3_RAT	Q61704	mus musculus	827	34.5	23.5	480	1	KRAC_MOUSE	P31750	mus musculus
755	35	23.8	887	1	ITH3_RAT	Q63416	rattus norv	828	34.5	23.5	480	1	KRAC_RAT	P47196	rattus norv
756	35	23.8	908	1	MGR8_HUMAN	O00222	homo sapien	829	34.5	23.5	492	1	CPA2_RAT	P15149	rattus norv
757	35	23.8	942	1	DP01_CHLAU	O08307	chloroflexu	830	34.5	23.5	501	1	KAKT_MLVAT	P31748	akt8 murine
758	35	23.8	977	1	EPA2_MOUSE	Q03145	mus musculus	831	34.5	23.5	510	1	ACHG_XENLA	P05376	xenopus lae
759	35	23.8	977	1	MCR_TUPGB	Q29131	tupaia glii	832	34.5	23.5	529	1	LAC1_PLEOS	Q12729	pleurotus o
760	35	23.8	981	1	MCR_RAT	P22199	rattus norv	833	34.5	23.5	533	1	INA6_MOUSE	Q35345	mus musculus
761	35	23.8	983	1	YQG4_CAEEL	Q09499	caenorhabdi	834	34.5	23.5	535	1	ICAI_BOVIN	Q95132	bos taurus
762	35	23.8	984	1	MCR_HUMAN	P08235	homo sapien	835	34.5	23.5	540	1	AXEL_ARATH	P42744	arabidopsis
763	35	23.8	1034	1	ENTK_PIG	P98074	sus scrofa	836	34.5	23.5	543	1	IEFS_HUMAN	P31948	homo sapien

837	34.5	23.5	553	1	TF3B_CANAL	P43072	candida alb	910	34	23.1	304	1	GDB1_WHEAT	P04729	tritium ae
838	34.5	23.5	608	1	ALBU_FELCA	P49064	felis silve	911	34	23.1	305	1	IBP2_MOUSE	P47877	mus musculus
839	34.5	23.5	676	1	KPCD_HUMAN	Q05655	homo sapien	912	34	23.1	306	1	YE16_YEAST	P40101	saccharomyc
840	34.5	23.5	695	1	TRFE_RABIT	P19134	oryctolagus	913	34	23.1	311	1	LECI_HUMAN	P07307	homo sapien
841	34.5	23.5	700	1	CAOQ_RAT	P19134	oryctolagus	914	34	23.1	313	1	YC85_SYNY3	P73191	synecocyst
842	34.5	23.5	781	1	YKG7_CAEEL	Q63448	rattus norv	915	34	23.1	323	1	FASA_BOVIN	P51867	bos taurus
843	34.5	23.5	831	1	PRIA_SYNY3	P46957	caenorhabdi	916	34	23.1	324	1	LEP_SALTY	P23697	salmonella
844	34.5	23.5	835	1	CD97_HUMAN	P74397	synecocyst	917	34	23.1	326	1	LACD_STAAU	P11100	staphylococ
845	34.5	23.5	842	1	LPFC_SALTY	P48960	homo sapien	918	34	23.1	328	1	IBP2_MOUSE	P18065	homo sapien
846	34.5	23.5	886	1	MCM2_XENLA	P43662	salmonella	919	34	23.1	340	1	PER_PERAM	P25637	periplaneta
847	34.5	23.5	900	1	SYM_HUMAN	P55861	xenopus lae	920	34	23.1	345	1	APDH_CANFA	P33703	canis famli
848	34.5	23.5	901	1	ODP1_MYCTU	P56192	homo sapien	921	34	23.1	350	1	PLEK_HUMAN	P08567	homo sapien
849	34.5	23.5	1046	1	YJ14_YEAST	Q10504	mycobacteri	922	34	23.1	356	1	VP39_NPVLD	P35840	lymantria d
850	34.5	23.5	1134	1	TIEL_MOUSE	P47029	saccharomyc	923	34	23.1	357	1	MTGL_HAEGA	P25282	haemophilus
851	34.5	23.5	1253	1	MYS9_DROME	Q06806	mus musculus	924	34	23.1	358	1	PIAP_PIG	O62640	sus scrofa
852	34.5	23.5	1335	1	RRPO_FXMV	Q01989	drosophila	925	34	23.1	360	1	AROC_ECOLI	P12008	escherichia
853	34.5	23.5	1389	1	PRAX_RAT	P22168	foxtail mos	926	34	23.1	366	1	CASA_EPHMU	P18503	ephydatia m
854	34.5	23.5	1390	1	MET_HUMAN	O63425	rattus norv	927	34	23.1	375	1	ADHH_GADMO	P81500	gadus morhu
855	34.5	23.5	1416	1	YN81_CAEEL	P08581	homo sapien	928	34	23.1	376	1	ADHL_GADMO	P81601	gadus morhu
856	34.5	23.5	3140	1	POLG_PPVSX	Q03610	caenorhabdi	929	34	23.1	382	1	ADHX_SPAU	P79896	sparus aur
857	34.5	23.5	4753	1	LRP_CAEEL	Q84934	p genome po	930	34	23.1	382	1	C79B_ARATH	O81346	arabidopsis
858	34	23.1	45	1	RUBR_DESDE	Q04833	caenorhabdi	931	34	23.1	383	1	LMX1_MESAU	Q04650	mesocricetu
859	34	23.1	62	1	OE18_NPVAC	P04170	desulfovibr	932	34	23.1	384	1	YGGW_HAEIN	P43899	haemophilus
860	34	23.1	69	1	RUBL_BRAJA	P48344	bradyrhizob	933	34	23.1	384	1	ADHX_CAEEL	Q17335	caenorhabdi
861	34	23.1	101	1	TSHE_PROSU	O62590	rhodopus su	934	34	23.1	384	1	GRPR_RAT	P52500	rattus norv
862	34	23.1	104	1	CHLE_PICMA	P37853	picea maria	935	34	23.1	388	1	FEM3_CAEEL	P34691	caenorhabdi
863	34	23.1	105	1	LSHB_MOUSE	Q09108	mus musculus	936	34	23.1	392	1	DDL_STRCO	Q92br9	streptomyce
864	34	23.1	111	1	PHNA_ECOLI	P16680	escherichia	937	34	23.1	392	1	YKH7_YEAST	P36081	saccharomyc
865	34	23.1	115	1	TR10_HUMAN	Q15642	homo sapien	938	34	23.1	393	1	METK_PINBN	P50300	pinus banks
866	34	23.1	127	1	ULB0_HCMVA	P16830	human cytom	939	34	23.1	393	1	OM1L_CHLTR	P19542	chlamydia t
867	34	23.1	136	1	EXB2_XANCP	Q34258	xanthomonas	940	34	23.1	394	1	OM1M_CHLTR	P06597	chlamydia t
868	34	23.1	138	1	TSHE_MOUSE	P12656	mus musculus	941	34	23.1	395	1	SDC_DROME	P49415	drosophila
869	34	23.1	138	1	TSHE_RAT	P04652	rattus norv	942	34	23.1	414	1	CP51_ISSOR	Q02315	issatchenk
870	34	23.1	141	1	LSHB_RAT	P01230	rattus norv	943	34	23.1	419	1	DNJH_ARATH	P42825	arabidopsis
871	34	23.1	142	1	RHIC_ORYRH	Q76145	oryctes rhi	944	34	23.1	421	1	EF1A_HALHA	P48863	halobacteri
872	34	23.1	148	1	Y219_MXCGE	P47461	mycoplasma	945	34	23.1	431	1	VMSA_WHV7	P12909	woodchuck h
873	34	23.1	149	1	PRO7_ORYSA	P20698	oryza sativ	946	34	23.1	431	1	VMSA_WHV8	P06432	woodchuck h
874	34	23.1	153	1	SODC_DEBHA	O42724	debaromyce	947	34	23.1	431	1	VMSA_WHV81	P17400	woodchuck h
875	34	23.1	159	1	YO21_CABEL	P34671	caenorhabdi	948	34	23.1	432	1	PYRP_BACCL	P41006	bacillus ca
876	34	23.1	160	1	VG34_HSVBE	P28989	equine herp	949	34	23.1	437	1	IF2G_METJA	Q38657	methanococc
877	34	23.1	167	1	RUVX_SYNY3	O55506	synecocyst	950	34	23.1	440	1	AK_CHLPN	Q92610	chlamydia p
878	34	23.1	169	1	LSHB_EQUBU	O46641	equus burch	951	34	23.1	444	1	OPCA_SYNP7	O54709	synecococc
879	34	23.1	169	1	LSHB_HORSE	P08751	equus cabal	952	34	23.1	444	1	RGS9_RAT	P49805	rattus norv
880	34	23.1	169	1	YRDD_ECOLI	P45771	escherichia	953	34	23.1	445	1	SIG_BORBU	O51344	borrelia bu
881	34	23.1	180	1	HYCF_ECOLI	P16432	escherichia	954	34	23.1	447	1	AC11_CAEEL	P34429	caenorhabdi
882	34	23.1	188	1	NEFB_ECOLI	P32707	escherichia	955	34	23.1	450	1	PX8A_HUMAN	Q06710	homo sapien
883	34	23.1	203	1	YEV8_YEAST	P40080	saccharomyc	956	34	23.1	457	1	PAX8_MOUSE	Q00288	mus musculus
884	34	23.1	204	1	MSX3_MOUSE	P70354	mus musculus	957	34	23.1	459	1	PX8A_CANFA	P47240	canis famli
885	34	23.1	213	1	NO2B_SOYBN	P08863	glycine max	958	34	23.1	462	1	NIFK_METMP	P71527	methanococc
886	34	23.1	220	1	PC83_SOLTU	P20347	solanum tub	959	34	23.1	463	1	VL2_HPV13	Q02275	human papil
887	34	23.1	220	1	PK1X_SOLTU	O00652	solanum tub	960	34	23.1	476	1	RF3_SACUV	P05512	saccharomyc
888	34	23.1	232	1	NIFO_RHTSN	Q53213	rhizobium s	961	34	23.1	482	1	CATA_VIBFI	O68146	vibrio fisc
889	34	23.1	237	1	P1V6_ADEM1	P48310	mouse adeno	962	34	23.1	492	1	CAT1_RICCO	Q01297	ricinus com
890	34	23.1	240	1	CD7_HUMAN	P09584	homo sapien	963	34	23.1	492	1	CAT2_RICCO	P49318	ricinus com
891	34	23.1	244	1	NDF3_MOUSE	P70660	mus musculus	964	34	23.1	492	1	CATA_HELAN	P45739	helianthus
892	34	23.1	249	1	HMX1_CHICK	P28361	gallus gall	965	34	23.1	492	1	CATA_IPOBA	P07145	ipomoea bat
893	34	23.1	250	1	LFA3_HUMAN	P19256	homo sapien	966	34	23.1	508	1	CATA_HAEIN	P44390	haemophilus
894	34	23.1	255	1	YQXN_BACSU	P42095	bacillus su	967	34	23.1	510	1	ERM_HUMAN	P41161	homo sapien
895	34	23.1	269	1	INHA_MYCSM	P42829	mycobacteri	968	34	23.1	512	1	HYA1_MOUSE	P48794	mus musculus
896	34	23.1	269	1	PENK_MYCSM	P22005	mus musculus	969	34	23.1	513	1	KAA8_SCHPO	Q05792	schizosacch
897	34	23.1	273	1	PRXC_PSEFL	O31158	pseudomonas	970	34	23.1	513	1	VAB1_BOVIN	P31407	bos taurus
898	34	23.1	275	1	RPB3_HUMAN	P19387	homo sapien	971	34	23.1	542	1	C79B_SINAL	O81345	sinapis alb
899	34	23.1	275	1	RPB3_MOUSE	P97760	mus musculus	972	34	23.1	550	1	IDS_HUMAN	P22304	homo sapien
900	34	23.1	279	1	ATND_RAT	Q63377	rattus norv	973	34	23.1	553	1	PRIS_DESVH	P31101	desulfovibr
901	34	23.1	281	1	VNS1_INBGA	P12592	influenza b	974	34	23.1	556	1	HEXB_HUMAN	P07686	homo sapien
902	34	23.1	281	1	VNS1_INBHK	P12594	influenza b	975	34	23.1	559	1	REF3_YEAST	P12294	saccharomyc
903	34	23.1	281	1	VNS1_INBID	P12596	influenza b	976	34	23.1	563	1	MUCS_BOVIN	P98091	bos taurus
904	34	23.1	283	1	GAGL_HUMAN	P10263	homo sapien	977	34	23.1	585	1	VMD2_HUMAN	O76090	homo sapien
905	34	23.1	284	1	SFAS_SPEST	P55925	spermatozop	978	34	23.1	586	1	LREL_YEAST	P25579	saccharomyc
906	34	23.1	287	1	CAPB_YEAST	P13517	saccharomyc	979	34	23.1	587	1	TALA_BFDV	P13894	budgerigar
907	34	23.1	288	1	YQD5_CAEEL	Q09265	caenorhabdi	980	34	23.1	587	1	YCAO_HAEIN	P41144	haemophilus
908	34	23.1	295	1	YFLF_CAUCR	P33976	caulobacter	981	34	23.1	588	1	CYDD_ECOLI	P29018	escherichia
909	34	23.1	301	1	NHAR_ECOLI	P10087	escherichia	982	34	23.1	591	1	CALX_MOUSE	P35564	mus musculus


```

983 34 23.1 591 1 CALX_RAT P35565 rattus norv
984 34 23.1 593 1 CALX_CANFA P24643 canis famil
985 34 23.1 600 1 IAP1_MOUSE O08863 mus musculu
986 34 23.1 604 1 IAP1_HUMAN Q13489 homo sapien
987 34 23.1 611 1 IAP_CHICK Q90660 gallus gall
988 34 23.1 611 1 MEN1_MOUSE O88559 mus musculu
989 34 23.1 612 1 IAP2_MOUSE Q62210 mus musculu
990 34 23.1 615 1 MEN1_HUMAN O00255 homo sapien
991 34 23.1 617 1 SG2_MOUSE Q03517 mus musculu
992 34 23.1 618 1 IAP2_HUMAN Q13490 homo sapien
993 34 23.1 626 1 PC11_YEAST P39081 saccharomyc
994 34 23.1 627 1 ACH4_HUMAN P43681 homo sapien
995 34 23.1 638 1 GHR_PIG P19756 sus scrofa
996 34 23.1 642 1 TRFE_GADMO Q92079 gadus morchu
997 34 23.1 650 1 SYM_HELPY P56127 helicobacte
998 34 23.1 658 1 KPCL_LYPTI Q25378 lytechinus
999 34 23.1 663 1 GLI3_CHICK P55879 gallus gall

```

ALIGNMENTS

```

RESULT 1
ID GLHA_MACMU STANDARD; PRT; 120 AA.
AC P22762;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GLYCOPROTEIN HORMONES ALPHA CHAIN PRECURSOR.
GN CGA.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopitheciinae;
OC Macaca.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91321740.
RA Golos T.G., Durning M., Fisher J.M.;
RT "Molecular cloning of the rhesus glycoprotein hormone alpha-subunit
gene.";
RL DNA Cell Biol. 10:367-380(1991).
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
FAMILY.
DR PIR; A39555; A39555.
DR HSP; P01215; IHRP.
DR PRINTS; PR00274; GLYCOHORMONE.
DR PRINTS; PR00438; GFCYSKNOT.
DR PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
DR PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
DR PFAM; PF00236; hormone6; 1.
KW Hormone; Glycoprotein; signal.
FT SIGNAL 1 24 BY SIMILARITY.
FT CHAIN 25 120 GLYCOPROTEIN HORMONES ALPHA CHAIN.
FT DISULFID 35 59 BY SIMILARITY.
FT DISULFID 38 88 BY SIMILARITY.
FT DISULFID 56 110 BY SIMILARITY.
FT DISULFID 60 112 BY SIMILARITY.
FT DISULFID 87 115 BY SIMILARITY.
FT CARBOHYD 67 67 BY SIMILARITY.
FT CARBOHYD 80 80 BY SIMILARITY.
FT CARBOHYD 106 106 BY SIMILARITY.
SQ SEQUENCE 120 AA; 13785 MW; 92E92D716093F406 CRC64;

```

Query Match 64.6%; Score 95; DB 1; Length 120;
 Best Local Similarity 64.38; Pred. No. 5.1e-07;
 Matches 18; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

```

Oy 2 pd-----vgdpectqlqenpfssqpgapi 25
   |||||:|||||:|||||:|||||
Db 26 PDGETMQDCPECKPRENKFESKPGAPI 53

RESULT 2
ID GLHA_MELGA STANDARD; PRT; 120 AA.
AC P37035;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GLYCOPROTEIN HORMONES ALPHA CHAIN PRECURSOR.
GN CGA.
OS Meleagris gallopavo (Common turkey), and
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC Neognathae; Galliformes; Meleagrididae; Meleagris.
RN [1]
RP SEQUENCE FROM N.A.
RX SPECIES-M.GALLOPAVO;
RX MEDLINE; 92150364.
RA Foster D.N., Foster L.K.;
RT "Cloning and sequence analysis of the common alpha-subunit
complementary deoxyribonucleic acid of turkey pituitary glycoprotein
hormones.";
RL Poult. Sci. 70:2516-2523(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX SPECIES-M.GALLOPAVO;
RX MEDLINE; 91097605.
RA Foster D.N., Kim S.U., Enyeart J.J., Foster L.K.;
RT "Nucleotide sequence of the complementary DNA for turkey growth
hormone.";
RL Biochem. Biophys. Res. Commun. 173:967-975(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX SPECIES-C.C.JAPONICA; TISSUE=PITUITARY;
RX MEDLINE; 94252550.
RA Ando H., Ishii S.;
RT "Molecular cloning of complementary deoxyribonucleic acids for the
pituitary glycoprotein hormone alpha-subunit and luteinizing hormone
beta-subunit precursor molecules of Japanese quail (Coturnix coturnix
japonica).";
RL Gen. Comp. Endocrinol. 93:357-368(1994).
RN [4]
RP SEQUENCE OF 34-61.
RX SPECIES-M.GALLOPAVO;
RX MEDLINE; 91065269.
RA Bergert E.R., Madden B., McCormick D.J., Pakkoff H., Ryan R.J.;
RT "The antigenic structure of the human glycoprotein hormone alpha-
subunit: II. Cross-species comparisons.";
RL Endocrinology 127:2985-2989(1990).
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M33698; AAA49629.1; -.
DR EMBL; S70833; AAB30866.1; -.
DR PIR; A45585; A45585.
DR HSP; P01215; IHRP.
DR PRINTS; PR00274; GLYCOHORMONE.

```


DR PRINTS; PR00438; GFCYSKNOT.
 DR PROSITE; PS00779; GLYCO.HORMONE_ALPHA_1; 1.
 DR PROSITE; PS00780; GLYCO.HORMONE_ALPHA_2; 1.
 DR PFAM; PF00236; hormone6; 1.
 KW Hormone; Glycoprotein; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 120 GLYCOPROTEIN HORMONES ALPHA CHAIN.
 FT DISULFID 35 59 BY SIMILARITY.
 FT DISULFID 38 88 BY SIMILARITY.
 FT DISULFID 56 110 BY SIMILARITY.
 FT DISULFID 60 112 BY SIMILARITY.
 FT DISULFID 87 115 BY SIMILARITY.
 FT CARBOHYD 80 80 POTENTIAL.
 FT CARBOHYD 106 106 POTENTIAL.
 SQ SEQUENCE 120 AA; 13591 MW; F4662BA0CB6005EB CRC64;

Query Match 61.9%; Score 91; DB 1; Length 120;
 Best Local Similarity 64.3%; Pred. No. 1.9e-06;
 Matches 18; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

QY 2 pd-----vqdceptlqenffsqpgapi 25
 || :| |||| | || ||||:||||
 DB 26 PDGEFLMQGCECKLGENRFFSKPGAPI 53

RESULT 3
 GLHA_STRCA STANDARD; PRT; 96 AA.
 AC P80665;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GLYCOPROTEIN HORMONES ALPHA CHAIN.
 GN CGA.
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
 OC Palaeognathae; Struthioniformes; Struthionidae; Struthio.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE; 97025333.
 RA Koide Y., Papkoff H., Kawachi H.;
 RT "Complete amino acid sequences of follitropin and lutropin in the ostrich, Struthio camelus.";
 RL Eur. J. Biochem. 240:262-267(1996).
 CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN, LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN FAMILY.

CC HSSP; P01215; 1HRP.
 DR PRINTS; PR00274; GLYCOHORMONE.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PROSITE; PS00779; GLYCO.HORMONE_ALPHA_1; 1.
 DR PROSITE; PS00780; GLYCO.HORMONE_ALPHA_2; 1.
 DR PFAM; PF00236; hormone6; 1.
 KW Hormone; Glycoprotein.
 FT DISULFID 11 35 BY SIMILARITY.
 FT DISULFID 14 64 BY SIMILARITY.
 FT DISULFID 32 86 BY SIMILARITY.
 FT DISULFID 36 88 BY SIMILARITY.
 FT DISULFID 63 91 BY SIMILARITY.
 FT CARBOHYD 56 56 POTENTIAL.
 FT CARBOHYD 82 82 POTENTIAL.
 SQ SEQUENCE 96 AA; 10781 MW; 1D6D4CC67187BC37 CRC64;

Query Match 61.2%; Score 90; DB 1; Length 96;
 Best Local Similarity 60.7%; Pred. No. 2.1e-06;
 Matches 17; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY 2 pd-----vqdceptlqenffsqpgapi 25

DB 2 PDGEFLMQGCECKLGENRFFSKPGAPV 29
 || :| |||| | || ||||:||||

RESULT 4
 GLHA_BOVIN STANDARD; PRT; 120 AA.
 ID GLHA_BOVIN
 AC P01217;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 23-OCT-1986 (Rel. 02, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GLYCOPROTEIN HORMONES ALPHA CHAIN PRECURSOR.
 GN CGA.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
 OC Bovinae; Bos.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 84041490.
 RA Goodwin R.G., Moncman C.L., Rottman F.M., Nilson J.H.;
 RT "Characterization and nucleotide sequence of the gene for the common alpha subunit of the bovine pituitary glycoprotein hormones.";
 RL Nucleic Acids Res. 11:6873-6882(1983).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 84024633.
 RA Erwin C., Croyle M.L., Donelson J., Maurer R.;
 RT "Nucleotide sequence of cloned complementary deoxyribonucleic acid for the alpha subunit of bovine pituitary glycoprotein hormones.";
 RL Biochemistry 22:4856-4860(1983).
 RN [3]
 RP SEQUENCE OF 8-133 FROM N.A.
 RX MEDLINE; 83161058.
 RA Nilson J.H., Thomason A.R., Cserbak M.T., Moncman C.L., Woychik R.P.;
 RT "Nucleotide sequence of a cDNA for the common alpha subunit of the bovine pituitary glycoprotein hormones. Conservation of nucleotides in the 3'-untranslated region of bovine and human pre-alpha subunit mRNAs.";
 RL J. Biol. Chem. 258:4679-4682(1983).
 RN [4]
 RP SEQUENCE OF 25-120.
 RX MEDLINE; 71111428.
 RA Liao T.-H., Pierce J.G.;
 RT "The primary structure of bovine thyrotropin. II. The amino acid sequences of the reduced, S-carboxymethyl alpha and beta chains.";
 RL J. Biol. Chem. 246:865-865(1971).
 RN [5]
 RP PRELIMINARY PARTIAL SEQUENCE.
 RX MEDLINE; 71111429.
 RA Pierce J.G., Liao T.-H., Carlsen R.B., Reimo T.;
 RT "Comparisons between the alpha chain of bovine thyrotropin and the C1 chain of luteinizing hormone. Compositions of tryptic peptides, cyanogen bromide fragments, and carbohydrate moieties.";
 RL J. Biol. Chem. 246:866-872(1971).
 RN [6]
 RP SEQUENCE OF 80-91 AND 100-120.
 RX MEDLINE; 72012932.
 RA Maguain-Rogister G., Hennen G.P.;
 RT "Bovine luteinizing hormone. Study of the primary structure around the carbohydrate attachment sites of the luteinizing hormone alpha-subunit.";
 RL Eur. J. Biochem. 21:489-497(1971).
 RN [7]
 RP PRELIMINARY ASSIGNMENT OF DISULFIDE BONDS.
 RX MEDLINE; 74307752.
 RA Corneli J.S., Pierce J.G.;
 RT "Studies on the disulfide bonds of glycoprotein hormones. Locations in the alpha chain based on partial reductions and formation of 14C-labeled S-carboxymethyl derivatives.";
 RL J. Biol. Chem. 249:4166-4174(1974).
 CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,


```
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
CC FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X00003; CAA24907.1; -.
DR EMBL; X00004; CAA24907.1; JOINED.
DR EMBL; X00050; CAA24932.1; -.
DR PIR; A01483; TTBOA.
DR PIR; A01532; A05132.
DR HSSP; P01215; IHRP.
DR PRINTS; PR00274; GLYCOHORMONE.
DR PRINTS; PR00438; GFCYSKNOT.
DR PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
DR PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
DR PFAM; PF00236; hormone6; 1.
KW Hormone; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 120 GLYCOPROTEIN HORMONES ALPHA CHAIN.
FT DISULFID 35 59 BY SIMILARITY.
FT DISULFID 38 88 BY SIMILARITY.
FT DISULFID 56 110 BY SIMILARITY.
FT DISULFID 60 112 BY SIMILARITY.
FT DISULFID 87 115 BY SIMILARITY.
FT CARBOHYD 80 80
FT CARBOHYD 106 106
FT CONFLICT 13 13 A -> T (IN REF. 2).
FT CONFLICT 37 37 E -> Q (IN REF. 4).
FT CONFLICT 52 52 P -> A (IN REF. 2).
FT CONFLICT 84 87 EATC -> AZCT (IN REF. 6).
SQ SEQUENCE 120 AA; 13616 MW; 2465415F86D04230 CRC64;

Query Match 57.8%; Score 85; DB 1; Length 120;
Best Local Similarity 57.1%; Pred. No. 1.4e-05;
Matches 16; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

QY 2 pd----vqdcpectlqenpfsgpgapi 25
   || ||||| ||| ||| |||
Db 26 PDGEFTMQGCECKLKENKYFKPDAP1 53

RESULT 5
GLHA_SHEEP STANDARD; PRT; 120 AA.
ID A01218;
AC P01218;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE GLYCOPROTEIN HORMONES ALPHA CHAIN PRECURSOR.
GN CGA.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
RN [1]
RX SEQUENCE FROM N.A.
RP MEDLINE; 90098887.
RA Bello P.A., Mountford P.S., Brandon M.R., Adams T.E.;
RT "Cloning and DNA sequence analysis of the cDNA for the common alpha-
RL subunit of the ovine pituitary glycoprotein hormones.";
RN Nucleic Acids Res. 17:10494-10494(1989).
RN [2]
RP SEQUENCE OF 25-120.
RX MEDLINE; 72211144.
```

```
RA Liu W.-K., Nahm H.S., Sweeney C.M., Lamkin W.M., Baker H.N.,
RA Ward D.N.;
RT "The primary structure of ovine luteinizing hormone. I. The amino
RT acid sequence of the reduced and S-aminoethylated S-subunit (LH-
RT alpha).";
RL J. Biol. Chem. 247:4351-4364(1972).
RL [3]
RN SEQUENCE OF 25-120.
RX MEDLINE; 73190034.
RA Sairam M.R., Papkoff H., Li C.H.;
RT "The primary structure of ovine interstitial cell-stimulating
RT hormone. I. The alpha-subunit.";
RL Arch. Biochem. Biophys. 153:554-571(1972).
RL [4]
RN SEQUENCE OF 25-120.
RX MEDLINE; 82113052.
RA Sairam M.R.;
RT "Primary structure of the ovine pituitary follitropin alpha-subunit.";
RL Biochem. J. 197:535-539(1981).
RL [5]
RN PRELIMINARY ASSIGNMENT OF DISULFIDE BONDS.
RA Chung D., Sairam M.R., Li C.H.;
RT "The primary structure of ovine interstitial cell-stimulating
RT hormone. III. Disulfide bridges of the alpha-subunit.";
RL Arch. Biochem. Biophys. 159:678-682(1973).
RL [6]
RN STRUCTURE OF CARBOHYDRATES.
RX MEDLINE; 91006170.
RA Weisshaar G., Hiyama J., Renwick A.G.;
RT "Site-specific N-glycosylation of ovine lutropin. Structural analysis
RT by one- and two-dimensional 1H-NMR spectroscopy.";
RL Eur. J. Biochem. 192:741-751(1990).
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
CC FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X16977; CAA34848.1; -.
DR PIR; A01483; TTBOA.
DR PIR; A01484; UTSHA.
DR PIR; S06935; S06935.
DR PIR; S13200; S13200.
DR HSSP; P01215; IHRP.
DR PRINTS; PR00274; GLYCOHORMONE.
DR PRINTS; PR00438; GFCYSKNOT.
DR PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
DR PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
DR PFAM; PF00236; hormone6; 1.
KW Hormone; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 120 GLYCOPROTEIN HORMONES ALPHA CHAIN.
FT DISULFID 35 59 BY SIMILARITY.
FT DISULFID 38 88 BY SIMILARITY.
FT DISULFID 56 110 BY SIMILARITY.
FT DISULFID 60 112 BY SIMILARITY.
FT DISULFID 87 115 BY SIMILARITY.
FT CARBOHYD 80 80
FT CARBOHYD 106 106
FT CONFLICT 27 27
FT CONFLICT 29 29 E -> Q (IN REF. 2).
FT CONFLICT 37 37 E -> Q (IN REF. 2).
FT CONFLICT 112 113 CS -> SC (IN REF. 3).
SQ SEQUENCE 120 AA; 13588 MW; 1F06F784F2420181 CRC64;
```



```
DR PFAM; PF00236; hormone6; 1.
KW Hormone; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 120 GLYCOPROTEIN HORMONES ALPHA CHAIN 1.
FT DISULFID 35 59 BY SIMILARITY.
FT DISULFID 38 88 BY SIMILARITY.
FT DISULFID 56 110 BY SIMILARITY.
FT DISULFID 60 112 BY SIMILARITY.
FT DISULFID 87 115 BY SIMILARITY.
FT CARBOHYD 80 80 PROBABLE.
FT CARBOHYD 106 106 PROBABLE.
FT CONFLICT 84 84 E -> Q (IN AAA97425).
SQ SEQUENCE 120 AA; 13453 MW; BE0E7F3C44C69ECB CRC64;

Query Match 57.8%; Score 85; DB 1; Length 120;
Best Local Similarity 57.1%; Pred. No. 1.4e-05;
Matches 16; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

Oy 2 pd----vqdcpectlqenpfsspgapi 25
|| :||| :||| :||| :|||
Db 26 PDGETMQCPECKLKNYFKSPDAP1 53

RESULT 6
GLH1_RAT STANDARD; PRT; 120 AA.
ID GLH1_RAT
AC P11962; P70516;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GLYCOPROTEIN HORMONES ALPHA CHAIN 1 PRECURSOR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=PIUITARY;
RX MEDLINE; 82214055.
RA Godine J.E., Chin W.W., Habener J.F.;
RT "Alpha subunit of rat pituitary glycoprotein hormones. Primary
RT structure of the precursor determined from the nucleotide sequence of
RT cloned cDNAs.";
RL J. Biol. Chem. 257:8368-8371(1982).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89196918.
RA Burnside J., Buckland P.R., Chin W.W.;
RT "Isolation and characterization of the gene encoding the alpha-subunit
RT of the rat pituitary glycoprotein hormones.";
RL Gene 70:67-74 (1988).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR-IMAMICHI; TISSUE=ANTERIOR PITUITARY;
RA Kato Y., Ezashi T., Hirai T., Kato T.;
RT "Strain difference in nucleotide sequences of rat glycoprotein hormone
RT subunit cDNAs and gene fragment.";
RL Zool. Sci. 7:877-885(1990).
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01252; AAA24565.1; -
DR EMBL; J00757; AAA97425.1; -
DR EMBL; M22829; -; NOT ANNOTATED_CDS.
DR EMBL; M25543; AAB04659.1; -
DR EMBL; M25344; AAB04668.1; -
DR EMBL; D00575; BAA00453.1; -
DR PIR; JT0408; TTRTA.
DR PIR; S27385; S27385.
DR HSSP; P01215; LHRP.
DR PRINTS; PR00274; GLYCOHORMONE.
DR PRINTS; PR00438; GFCYSKNOT.
DR PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
DR PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.

Oy 2 pd----vqdcpectlqenpfsspgapi 25
|| :||| :||| :||| :|||
Db 26 PDGDLTIQGCPECKLKNYFKSLGAP1 53

RESULT 7
GLH2_RAT STANDARD; PRT; 120 AA.
ID GLH2_RAT
AC P11963;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GLYCOPROTEIN HORMONES ALPHA CHAIN 2 PRECURSOR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 82214055.
RA Godine J.E., Chin W.W., Habener J.F.;
RT "Alpha subunit of rat pituitary glycoprotein hormones. Primary
RT structure of the precursor determined from the nucleotide sequence of
RT cloned cDNAs.";
RL J. Biol. Chem. 257:8368-8371(1982).
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01253; CAA24566.1; -
DR PIR; S27386; S27386.
DR HSSP; P01215; LHRP.
DR PRINTS; PR00274; GLYCOHORMONE.
DR PRINTS; PR00438; GFCYSKNOT.
DR PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
DR PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
DR PFAM; PF00236; hormone6; 1.
KW Hormone; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 120 GLYCOPROTEIN HORMONES ALPHA CHAIN 2.
FT DISULFID 35 59 BY SIMILARITY.
FT DISULFID 38 88 BY SIMILARITY.
FT DISULFID 56 110 BY SIMILARITY.
FT DISULFID 60 112 BY SIMILARITY.
```


FT DISULFID 87 115 BY SIMILARITY.
 FT CARBOHYD 80 80 PROBABLE.
 FT CARBOHYD 106 106 PROBABLE.
 SQ SEQUENCE 120 AA; 13565 MW; 1BE324C29737B6D1 CRC64;

Query Match 57.1%; Score 84; DB 1; Length 120;
 Best Local Similarity 57.1%; Pred. No. 1.9e-05;
 Matches 16; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

OY 2 pd----vqdcpectqlqenffsqgapi 25
 || : ||||| : || : || : |||||
 Db 26 PDGDFIIQGCPECKLKENKYSKLGAPI 53

RESULT 8

GLHA_MOUSE
 ID GLHA_MOUSE STANDARD; PRT; 120 AA.
 AC P01216;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 13-DEC-1998 (Rel. 37, Last annotation update)
 DE GLYCOPROTEIN HORMONES ALPHA CHAIN PRECURSOR.
 CGA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 82060239.
 RA Chin W.W., Kronenberg H.M., Dee P.C., Maloof F., Habener J.F.;
 RT "Nucleotide sequence of the mRNA encoding the pre-alpha-subunit of
 mouse thyrotropin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 78:5329-5333(1981).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 89170115.
 RA Gordon D.F., Wood W.M., Ridgway E.C.;
 RT "Organization and nucleotide sequence of the mouse alpha-subunit gene
 of the pituitary glycoprotein hormones.";
 RL DNA 7:679-690(1988).
 CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
 CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
 LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
 FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J00643; AAA96700.1; -;
 DR EMBL; V00852; CAA24223.1; -;
 DR EMBL; M22991; -; NOT_ANNOTATED_CDS.
 DR EMBL; M22992; AAA99228.1; -;
 DR PIR; A01482; TTWSA.
 DR PIR; A31598; A31598.
 DR HSSP; P01215; 1HRP.
 DR MGD; MGI:88390; CGA.
 DR PRINTS; PR00274; GLYCOPROTEIN.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PROSITE; PS00779; GLYCOPROTEIN_HORMONE_ALPHA_1; 1.
 DR PROSITE; PS00780; GLYCOPROTEIN_HORMONE_ALPHA_2; 1.
 DR PFAM; PF00236; hormone6; 1.
 KW Hormone; Glycoprotein; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 120 GLYCOPROTEIN HORMONES ALPHA CHAIN.
 FT DISULFID 35 59 BY SIMILARITY.

FT DISULFID 38 88 BY SIMILARITY.
 FT DISULFID 56 110 BY SIMILARITY.
 FT DISULFID 60 112 BY SIMILARITY.
 FT DISULFID 87 115 BY SIMILARITY.
 FT CARBOHYD 80 80 PROBABLE.
 FT CARBOHYD 106 106 PROBABLE.
 SQ SEQUENCE 120 AA; 13565 MW; F2FA2E4632716DBC CRC64;

Query Match 57.1%; Score 84; DB 1; Length 120;
 Best Local Similarity 57.1%; Pred. No. 1.9e-05;
 Matches 16; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

OY 2 pd----vqdcpectqlqenffsqgapi 25
 || : ||||| : || : || : |||||
 Db 26 PDGDFIIQGCPECKLKENKYSKLGAPI 53

RESULT 9

GLHA_RABIT
 ID GLHA_RABIT STANDARD; PRT; 96 AA.
 AC P07474;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GLYCOPROTEIN HORMONES ALPHA CHAIN.
 CGA.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 RN [1]
 RP SEQUENCE.
 RA Glenn S.D., Nahm H.S., Ward D.N.;
 RT "The amino acid sequence of the rabbit glycoprotein hormone alpha
 subunit.";
 RL J. Protein Chem. 3:143-156(1984).
 CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
 CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
 LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
 FAMILY.

PIR; A05096; A05096.
 DR HSSP; P01215; 1HRP.
 DR PRINTS; PR00274; GLYCOPROTEIN.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PROSITE; PS00779; GLYCOPROTEIN_HORMONE_ALPHA_1; 1.
 DR PROSITE; PS00780; GLYCOPROTEIN_HORMONE_ALPHA_2; 1.
 DR PFAM; PF00236; hormone6; 1.
 KW Hormone; Glycoprotein.
 FT DISULFID 11 35 BY SIMILARITY.
 FT DISULFID 14 64 BY SIMILARITY.
 FT DISULFID 32 86 BY SIMILARITY.
 FT DISULFID 36 88 BY SIMILARITY.
 FT DISULFID 63 91 BY SIMILARITY.
 FT CARBOHYD 56 56
 FT CARBOHYD 82 82
 SQ SEQUENCE 96 AA; 10663 MW; 0CB16491940F6300 CRC64;

Query Match 55.8%; Score 82; DB 1; Length 96;
 Best Local Similarity 57.1%; Pred. No. 3e-05;
 Matches 16; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

OY 2 pd----vqdcpectqlqenffsqgapi 25
 || : ||||| : || : || : |||||
 Db 2 PDGEFAMQGCPECKLKENKYSKLGAPI 29

RESULT 10

GLHA_MACRU
 ID GLHA_MACRU STANDARD; PRT; 120 AA.
 AC O46687;


```
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE GLYCOPROTEIN HORMONES ALPHA CHAIN PRECURSOR.
GN CCA.
OS Macropus rufus (Red kangaroo) (Megaleia rufa), and
OS Trichosurus vulpecula (Brush-tailed possum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Metatheria; Diprotodontia; Macropodidae; Macropus.
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=PIUITARY;
RX MEDLINE; 98345424.
RA Harrison G.A., Deane E.M., Cooper D.W.;
RT "cDNA cloning of luteinizing hormone subunits from brushtail possum
and red kangaroo.";
RL Mamm. Genome 9:638-642(1998).
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF017449; AAC96020.1; -.
DR EMBL; AF017447; AAC96018.1; -.
DR PRINTS; PR00274; GLYCOHORMONE.
DR PRINTS; PR00438; GFCYSKNOT.
DR PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
DR PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
DR PFAM; PF00236; hormone6; 1.
KW Hormone; Glycoprotein; Signal.
FT SIGNAL 1 24 BY SIMILARITY.
FT CHAIN 25 120 GLYCOPROTEIN HORMONES ALPHA CHAIN.
FT DISULFID 35 59 BY SIMILARITY.
FT DISULFID 38 88 BY SIMILARITY.
FT DISULFID 56 110 BY SIMILARITY.
FT DISULFID 60 112 BY SIMILARITY.
FT DISULFID 87 115 BY SIMILARITY.
FT CARBOHYD 80 80 POTENTIAL.
FT CARBOHYD 106 106 POTENTIAL.
SQ SEQUENCE 120 AA; 13513 MW; 837F323310E280CB CRC64;

Query Match 55.8%; Score 82; DB 1; Length 120;
Best Local Similarity 57.1%; Pred. No. 3.8e-05;
Matches 16; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

QY 2 pd-----vqdcpectlqenpfsgqapi 25
|| : ||||| :||| :|||
DB 26 PDGEFINGQCPCKLKENKYFSLGAPI 53

RESULT 11
GLHA_PIG
ID GLHA_PIG STANDARD; PRT; 120 AA.
AC P01219;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GLYCOPROTEIN HORMONES ALPHA CHAIN PRECURSOR.
GN CGA.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
```

```
RN SEQUENCE FROM N.A.
RX MEDLINE; 89325834.
RA Hirai T., Takikawa H., Kato Y.;
RT "Molecular cloning of cDNAs for precursors of porcine pituitary
glycoprotein hormone common alpha-subunit and of thyroid stimulating
hormone beta-subunit.";
RL Mol. Cell. Endocrinol. 63:209-217(1989).
RN [2]
RP SEQUENCE FROM N.A.
RA Kato Y., Ezashi T., Hirai T., Kato T.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 31-120.
RX MEDLINE; 74075725.
RA Maghain-Rogister G., Combarnous Y., Hennen G.;
RT "The primary structure of the porcine luteinizing-hormone
alpha-subunit.";
RL Eur. J. Biochem. 39:255-263(1973).
RN [4]
RP PRELIMINARY SEQUENCE OF 25-120.
RA Closset J., Maghain-Rogister G., Hennen G.;
RL Endocrinol. Exp. 8:164-164(1974).
RN [5]
RP PRELIMINARY ASSIGNMENT OF DISULFIDE BONDS.
RX MEDLINE; 75093922.
RA Combarnous Y., Hennen G.;
RT "The disulphide bridges of porcine luteinizing hormone alpha
subunit.";
RL Biochem. Soc. Trans. 2:915-917(1974).
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D00768; BAA00664.1; -.
DR EMBL; D00767; BAA00664.1; JOINED.
DR PIR; A01485; UTEGA.
DR PIR; A30339; A30339.
DR HSP; P01215; IHRP.
DR PRINTS; PR00274; GLYCOHORMONE.
DR PRINTS; PR00438; GFCYSKNOT.
DR PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
DR PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
DR PFAM; PF00236; hormone6; 1.
KW Hormone; Glycoprotein; Signal.
FT SIGNAL 1 24 GLYCOPROTEIN HORMONES ALPHA CHAIN.
FT CHAIN 25 120 BY SIMILARITY.
FT DISULFID 35 59 BY SIMILARITY.
FT DISULFID 38 88 BY SIMILARITY.
FT DISULFID 56 110 BY SIMILARITY.
FT DISULFID 60 112 BY SIMILARITY.
FT DISULFID 87 115 BY SIMILARITY.
FT CARBOHYD 80 80
FT CARBOHYD 106 106
SQ SEQUENCE 120 AA; 13532 MW; 1DCBB2A375F192E6 CRC64;

Query Match 55.8%; Score 82; DB 1; Length 120;
Best Local Similarity 57.1%; Pred. No. 3.8e-05;
Matches 16; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

QY 2 pd-----vqdcpectlqenpfsgqapi 25
```



```

Db 26 PDGEFTMGCGECKLKNKYFSKLGAPI 53
      || :||| | :|| :|| :|||
      15-DEC-1998 (Rel. 37, Last annotation update)
      GLYCOPROTEIN HORMONES ALPHA CHAIN PRECURSOR.
CGA.
OS Callithrix jacchus (Common marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Platyrrhini; Callitrichidae; Callithrix.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PLACENTA;
RX MEDLINE; 96115012.
RA Simula A.P., Amato F., Faast R., Lopata A., Berka J., Norman R.J.;
RT "Luteinizing hormone/chorionic gonadotropin bioactivity in the common
RT marmoset (Callithrix jacchus) is due to a chorionic gonadotropin
RT molecule with a structure intermediate between human chorionic
RT gonadotropin and human luteinizing hormone.";
RL Biol. Reprod. 53:380-389(1995).
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U04446; AAC00030.1; -
CC HSP; P01215; IHRP.
CC PRINTS; PR00274; GLYCOHORMONE.
CC PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
CC PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
CC PFAM; PF00236; hormone6; 1.
CC Hormone; Glycoprotein; Signal.
KW SIGNAL 1 24 BY SIMILARITY.
FT CHAIN 25 120 GLYCOPROTEIN HORMONES ALPHA CHAIN.
FT DISULFID 35 59 BY SIMILARITY.
FT DISULFID 38 88 BY SIMILARITY.
FT DISULFID 56 110 BY SIMILARITY.
FT DISULFID 60 112 BY SIMILARITY.
FT DISULFID 87 115 BY SIMILARITY.
FT CARBOHYD 67 67 BY SIMILARITY.
FT CARBOHYD 80 80 BY SIMILARITY.
FT CARBOHYD 106 106 BY SIMILARITY.
SQ SEQUENCE 120 AA; 13744 MW; 6EEF0638179D56F9 CRC64;

Query Match 53.1%; Score 78; DB 1; Length 120;
Best Local Similarity 50.0%; Pred. No. 0.00014;
Matches 14; Conservative 6; Mismatches 4; Indels 4; Gaps 1;

Qy 2 pd----vqdcpectlqenpfsgqapi 25
|| :||| | :|| :|| :|||
Db 26 PDGEFTMGCGECKLKNKYFSKLGAPI 53
      || :||| | :|| :|| :|||

RESULT 14
ID GLHA_HORSE
AC P01220;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GLYCOPROTEIN HORMONES ALPHA CHAIN PRECURSOR.
GN CGA.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Perissodactyla; Equidae; Equus.

Db 26 PDGEFTMGCGECKLKNKYFSKLGAPI 53
      || :||| | :|| :|| :|||
      15-DEC-1998 (Rel. 37, Last annotation update)
      GLYCOPROTEIN HORMONES ALPHA CHAIN PRECURSOR.
CGA.
OS Callithrix jacchus (Common marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Platyrrhini; Callitrichidae; Callithrix.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PLACENTA;
RX MEDLINE; 96115012.
RA Simula A.P., Amato F., Faast R., Lopata A., Berka J., Norman R.J.;
RT "Luteinizing hormone/chorionic gonadotropin bioactivity in the common
RT marmoset (Callithrix jacchus) is due to a chorionic gonadotropin
RT molecule with a structure intermediate between human chorionic
RT gonadotropin and human luteinizing hormone.";
RL Biol. Reprod. 53:380-389(1995).
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X85170; CAA59454.1; -
CC HSP; P01215; IHRP.
CC PRINTS; PR00274; GLYCOHORMONE.
CC PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
CC PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
CC PFAM; PF00236; hormone6; 1.
CC Hormone; Glycoprotein; Signal.
KW SIGNAL 1 24 BY SIMILARITY.
FT CHAIN 25 120 GLYCOPROTEIN HORMONES ALPHA CHAIN.
FT DISULFID 35 59 BY SIMILARITY.
FT DISULFID 38 88 BY SIMILARITY.
FT DISULFID 56 110 BY SIMILARITY.
FT DISULFID 60 112 BY SIMILARITY.
FT DISULFID 87 115 BY SIMILARITY.
FT CARBOHYD 67 67 BY SIMILARITY.
FT CARBOHYD 80 80 BY SIMILARITY.
FT CARBOHYD 106 106 BY SIMILARITY.
SQ SEQUENCE 120 AA; 13721 MW; 03B1C5A3B945B3DE CRC64;

Query Match 54.4%; Score 80; DB 1; Length 120;
Best Local Similarity 53.6%; Pred. No. 7.3e-05;
Matches 15; Conservative 4; Mismatches 5; Indels 4; Gaps 1;

Qy 2 pd----vqdcpectlqenpfsgqapi 25
|| :||| | :|| :|| :|||
Db 26 PDGEFTMGCGECKLKNKYFSKLGAPI 53
      || :||| | :|| :|| :|||

RESULT 13
ID GLHA_CALJA
AC P51499;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
```


DR PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.

FT	DISULFID	59	110	BY SIMILARITY.
----	----------	----	-----	----------------


```

FT DISULFID      86      113      BY SIMILARITY.
FT CARBOHYD      79      79      PROBABLE.
FT CARBOHYD     104     104      PROBABLE.
SQ SEQUENCE     118 AA; 13548 MW;  DFB590F899EFCF1E CRC64;

Query Match      49.7%; Score 73; DB 1; Length 118;
Best Local Similarity 63.2%; Pred. No. 0.00073;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      7      cpecltqenffsqpgapi 25
      |||||:|||||
DB      34      CEECKLKENNIFSRGAPV 52

RESULT 21
GLHA_CTEID
ID      GLHA_CTEID      STANDARD;      PRT;      118 AA.
AC      P30983;
DT      01-JUL-1993 (Rel. 26, Created)
DT      01-JUL-1993 (Rel. 26, Last sequence update)
DT      15-DEC-1998 (Rel. 37, Last annotation update)
DE      DE GLYCOPROTEIN HORMONES ALPHA CHAIN PRECURSOR (GONADOTROPIN ALPHA
DE      CHAIN) (GTH-ALPHA)
OS      Ctenopharyngodon idella (Grass carp).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC      Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC      Cyprinoidae; Cyprinidae; Cyprininae; Ctenopharyngodon.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=PIUTARY;
RC      Chang Y.S., Huang F.-L., Lo T.-B.;
RL      Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: INVOLVED IN GAMETOGENESIS AND STEROIDOGENESIS.
CC      -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC      -1- SUBCELLULAR LOCATION: SECRETED.
CC      -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
CC      FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL; X61050; CAA43384.1; --
DR      PIR; S16762; S16762.
DR      HSP; P01215; IHRP.
DR      PRINTS; PR00274; GLYCOPROTEIN.
DR      PRINTS; PR00438; GFCYSKNOT.
DR      PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
DR      PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
DR      PFAM; PF00236; hormone6; 1.
DR      Hormone; Glycoprotein; Signal.
DR      SIGNAL      1      23      BY SIMILARITY.
FT      CHAIN      24      118      GLYCOPROTEIN HORMONES ALPHA CHAIN.
FT      DISULFID      34      58      BY SIMILARITY.
FT      DISULFID      37      87      BY SIMILARITY.
FT      DISULFID      55      108      BY SIMILARITY.
FT      DISULFID      59      110      BY SIMILARITY.
FT      DISULFID      86      113      BY SIMILARITY.
FT      CARBOHYD      79      79      PROBABLE.
FT      CARBOHYD     104     104      PROBABLE.
SQ      SEQUENCE     118 AA; 13453 MW;  54F76D7B3A0470BB CRC64;

Query Match      49.7%; Score 73; DB 1; Length 118;
Best Local Similarity 63.2%; Pred. No. 0.00073;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      7      cpecltqenffsqpgapi 25

```



```
RC TISSUE-PITUITARY;
RX MEDLINE; 90092087.
RA Liu C.-S., Huang F.-L., Chang C.-H., Lo T.-B.;
RT "Pike eel (Muraenesox cinereus) gonadotropin. Amino acid sequences of
RL both alpha and beta subunits.";
RL Eur. J. Biochem. 186:105-114(1989).
CC -!- FUNCTION: INVOLVED IN GAMETOGENESIS AND STEROIDOGENESIS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
CC FAMILY.
DR PIR; S07091; S07091.
DR HSP; P01215; IHRP.
DR PRINTS; PR00274; GLYCOHORMONE.
DR PRINTS; PR00438; GFCYSKNOT.
DR PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
DR PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
DR PFAM; PF00236; hormone6; 1.
KW Hormone; Glycoprotein.
FT DISULFID 11 35 BY SIMILARITY.
FT DISULFID 14 64 BY SIMILARITY.
FT DISULFID 32 83 BY SIMILARITY.
FT DISULFID 36 85 BY SIMILARITY.
FT DISULFID 63 88 BY SIMILARITY.
FT CARBOHYD 79 79
SQ SEQUENCE 93 AA; 10563 MW; 2C5B3005B260F40C CRC64;

Query Match 49.0%; Score 72; DB 1; Length 93;
Best Local Similarity 63.2%; Pred. No. 0.00079;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 7 cpecltqlgenpfqsgapi 25
Db 11 CDECKLKDKNKFKSPAPI 29

RESULT 24
GLHI_ONCKE
ID GLHI_ONCKE STANDARD; PRT; 108 AA.
AC P13152;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 13-DEC-1998 (Rel. 37, Last annotation update)
DE GLYCOPROTEIN HORMONES ALPHA CHAIN 1 PRECURSOR (GONADOTROPIN 1 ALPHA
DE CHAIN) (GTH-ALPHA) (FRAGMENT).
OS Oncomorhynchus keta (Chum salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
OC Salmoniformes; Salmonidae; Oncorhynchus.
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89169356.
RA Kitahara N., Nishizawa T., Gatanaga T., Okazaki H., Andoh T.,
RA Soma G.-I.;
RT "Primary structure of two mRNAs encoding putative salmon
RT alpha-subunits of pituitary glycoprotein hormone.";
RL Comp. Biochem. Physiol. 91B:551-556(1988).
CC -!- FUNCTION: INVOLVED IN GAMETOGENESIS AND STEROIDOGENESIS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M27652; AAA49404.1; -.
```

```
DR PIR; JLO068; JLO068.
DR HSP; P01215; IHRP.
DR PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
DR PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
DR PFAM; PF00236; hormone6; 1.
KW Hormone; Glycoprotein; Signal.
FT NON_TER 1 1
FT SIGNAL <1 13
FT CHAIN 14 108 GLYCOPROTEIN HORMONES ALPHA CHAIN 1.
FT DISULFID 24 48 BY SIMILARITY.
FT DISULFID 27 77 BY SIMILARITY.
FT DISULFID 45 98 BY SIMILARITY.
FT DISULFID 49 100 BY SIMILARITY.
FT DISULFID 76 103 BY SIMILARITY.
FT CARBOHYD 69 69 POTENTIAL.
FT CARBOHYD 94 94
SQ SEQUENCE 108 AA; 12047 MW; 6F3194B27A419234 CRC64;

Query Match 49.0%; Score 72; DB 1; Length 108;
Best Local Similarity 63.2%; Pred. No. 0.00092;
Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 7 cpecltqlgenpfqsgapi 25
Db 24 CEECKLKENKVFNSPGAPV 42

RESULT 25
GLHA_RANCA
ID GLHA_RANCA STANDARD; PRT; 97 AA.
AC P80051;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GLYCOPROTEIN HORMONES ALPHA CHAIN
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Raninae; Rana.
[1]
RP SEQUENCE.
RX MEDLINE; 92111564.
RA Hayashi H., Hayashi T., Hanaoka Y.;
RT "Amphibian lutropin and follitropin from the bullfrog Rana
RT catesbeiana. Complete amino acid sequence of the alpha subunit.";
RL Eur. J. Biochem. 203:185-191(1992).
[2]
RP SEQUENCE OF 1-28.
RX MEDLINE; 9105289.
RA Bergert E.R., Madden B., McCormick D.J., Papkoff H., Ryan R.J.;
RT "The antigenic structure of the human glycoprotein hormone alpha-
RT subunit: II. Cross-species comparisons.";
RL Endocrinology 127:2985-2989(1990).
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
CC FAMILY.
DR PIR; S20287; S20287.
DR HSP; P01215; IHCN.
DR PRINTS; PR00274; GLYCOHORMONE.
DR PRINTS; PR00438; GFCYSKNOT.
DR PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
DR PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
DR PFAM; PF00236; hormone6; 1.
KW Hormone; Glycoprotein.
FT DISULFID 11 36
FT DISULFID 14 65 BY SIMILARITY.
FT DISULFID 33 87 BY SIMILARITY.
FT DISULFID 37 89 BY SIMILARITY.
FT CARBOHYD 64 92
FT CARBOHYD 57 57
```


FT CARBOHYD 83 83
SQ SEQUENCE 97 AA; 11036 MW; 1B7D72AA1773A107 CRC64;

Query Match 36.1%; Score 53; DB 1; Length 97;
Best Local Similarity 62.5%; Pred. No. 0.44;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 7 cpeclqlenpfssqpg 22
|||||:|||||

Db 11 CPECKLENLRFNSMG 26

RESULT 26

MRP3_RAT
ID MRP3_RAT STANDARD; PRT; 1522 AA.
AC 088563; 088270;
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE CANALICULAR MULTISPECIFIC ORGANIC ANION TRANSPORTER 2 (MULTIDRUG
DE RESISTANCE-ASSOCIATED PROTEIN 3) (MRP-LIKE PROTEIN-2) (MLP-2).
GN ABCC3 OR CMOAT2 OR MRP3 OR MLP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 99292429.
RA Ortiz D.F., Li S., Iyer R., Zhang X., Novikoff P., Arias I.M.;
RT "MRP3, a new ATP-binding cassette protein localized to the canalicular
RT domain of the hepatocyte";
RL Am. J. Physiol. 276:G1493-G1500(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=COLON;
RX MEDLINE; 98279126.
RA Hirohashi T., Suzuki H., Ito K., Ogawa K., Kume K., Shimizu T.,
RA Sugiyama Y.;
RT "Hepatic expression of multidrug resistance-associated protein-like
RT proteins maintained in eisa hyperbilirubinemic rats";
RL Mol. Pharmacol. 53:1068-1075(1998).
CC -!- FUNCTION: MAY ACT AS AN INDUCIBLE TRANSPORTER IN THE BILIARY AND
CC -!- INTESTINAL EXCRETION OF ORGANIC ANIONS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: LONG AND INTESTINE, LOW IN LIVER. HIGHER IN
CC LIVER OF EISA HYPERBILIRUBINEMIC RATS AND TR(-) MUTANT RATS.
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MRP SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF072816; AAC25416.1; -
DR EMBL; AB010467; BAA28955.1; -
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
DR HSSP; P13569; INBD.
DR PFAM; PF00005; ABC_tran; 2.
DR PFAM; PF00664; ABC_membrane; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat.
FT DOMAIN 1 32 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 33 53 1 (BY SIMILARITY).
FT DOMAIN 54 73 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 74 94 2 (BY SIMILARITY).
FT DOMAIN 95 99 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 100 120 3 (BY SIMILARITY).
FT DOMAIN 121 132 CYTOPLASMIC (BY SIMILARITY).

133 TRANSMEM 133
154 DOMAIN 154
171 TRANSMEM 171
192 TRANSMEM 192
301 DOMAIN 301
322 TRANSMEM 322
347 DOMAIN 347
368 TRANSMEM 368
424 DOMAIN 424
445 TRANSMEM 445
448 TRANSMEM 448
469 TRANSMEM 469
531 DOMAIN 531
552 TRANSMEM 552
574 DOMAIN 574
595 TRANSMEM 595
598 TRANSMEM 598
979 TRANSMEM 979
980 TRANSMEM 980
1016 TRANSMEM 1016
1037 TRANSMEM 1037
1080 TRANSMEM 1080
1101 TRANSMEM 1101
1102 TRANSMEM 1102
1123 TRANSMEM 1123
1194 TRANSMEM 1194
1215 TRANSMEM 1215
1217 TRANSMEM 1217
1238 TRANSMEM 1238
1522 TRANSMEM 1522
666 TRANSMEM 666
1325 TRANSMEM 1325
18 TRANSMEM 18
1001 TRANSMEM 1001
1002 TRANSMEM 1002
344 TRANSMEM 344
645 TRANSMEM 645
1075 TRANSMEM 1075
1522 AA; 168977 MW; 740E31E0C4C64297 CRC64;
SQ SEQUENCE 1522 AA; 168977 MW; 740E31E0C4C64297 CRC64;

Query Match 35.4%; Score 52; DB 1; Length 1522;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 10 ctlqlenpfssqpg 24
|||||:|||||

Db 701 CTLQENVLFQGMNP 715

RESULT 27

FINC_CANFA
ID FINC_CANFA STANDARD; PRT; 522 AA.
AC 028275; Q28276;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FIBRONECTIN (FN) (FRAGMENT).
GN FN1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT (V+C)-.
RC TISSUE=CARILAGE;
RX MEDLINE; 96324983.
RA MacLeod J.N., Burton-Wurster N., Gu D.N., Lust G.;
RT "Fibronectin mRNA splice variant in articular cartilage lacks bases
RT encoding the V, III-15, and I-10 protein segments";
RL J. Biol. Chem. 271:18954-18960(1996).
CC -!- FUNCTION: FIBRONECTINS BIND CELL SURFACES AND VARIOUS COMPOUNDS
CC INCLUDING COLLAGEN, FIBRIN, HEPARIN, DNA, AND ACTIN. FIBRONECTINS
CC ARE INVOLVED IN CELL ADHESION, CELL MOTILITY, OPSONIZATION, WOUND
CC HEALING, AND MAINTENANCE OF CELL SHAPE.

CC -!- FUNCTION: VARIANT (V+C)- IS PROBABLY INVOLVED IN MATRIX
 CC ORGANIZATION OF CARTILAGE.
 CC -!- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED
 CC VARIANTS, CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS;
 CC TO A LESSER EXTEND HOMODIMERS.
 CC -!- ALTERNATIVE PRODUCTS: EACH OF THE "EXTRA DOMAIN" & THE CONNECTING
 CC STRAND 3 ARE PRESENT IN SOME FORMS OF FIBRONECTIN AND ABSENT IN
 CC OTHERS. THESE DIFFERENCES ARE DUE TO ALTERNATIVE SPLICING. VARIANT
 CC (V+C)- LACKS REPEAT 15 OF FIBRONECTIN TYPE-III, REPEAT 10 OF
 CC FIBRONECTIN TYPE-I, AND THE CONNECTING STRAND 3.
 CC -!- TISSUE SPECIFICITY: PLASMA FN (SOLUBLE DIMERIC FORM) IS SECRETED
 CC BY HEPATOCYTES. CELLULAR FN (DIMERIC OR CROSS-LINKED MULTIMERIC
 CC FORMS), MADE BY FIBROBLASTS, EPITHELIAL AND OTHER CELL TYPES, IS
 CC DEPOSITED AS FIBRILS IN THE EXTRACELLULAR MATRIX. VARIANT (V+C)-
 CC IS THE MAJOR TRANSCRIPT IN ARTICULAR CARTILAGE, BUT IT IS ABSENT
 CC FROM LIVER.
 CC -!- SIMILARITY: CONTAINS AT LEAST 4 FIBRONECTIN TYPE III DOMAINS.
 CC -!- SIMILARITY: CONTAINS AT LEAST 3 FIBRONECTIN TYPE I DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U52106; AAC48612.1; -;
 CC EMBL; U52105; AAC48611.1; -;
 CC HSP; P02751; 1FBR.
 CC DR PROSITE; PS01253; FIBRONECTIN_1; 2.
 CC DR PFAM; PF00039; fn1; 2.
 CC DR PFAM; PF00041; fn3; 3.
 CC KW Glycoprotein; Plasma; Heparin-binding; Acute phase; Cell adhesion;
 CC Repeat; Alternative splicing.
 CC FT NON_TER 1 1
 CC FT DOMAIN <1 204 HEPARIN-BINDING 2 (BY SIMILARITY).
 CC FT DOMAIN 325 >522 FIBRIN-BINDING 2 (BY SIMILARITY).
 CC FT DOMAIN <1 25 FIBRONECTIN TYPE-III.
 CC FT DOMAIN 26 114 FIBRONECTIN TYPE-III.
 CC FT DOMAIN 115 204 FIBRONECTIN TYPE-III.
 CC FT DOMAIN 205 315 CONNECTING STRAND 3 (CS-3) (V REGION).
 CC FT DOMAIN 316 395 FIBRONECTIN TYPE-III.
 CC FT DOMAIN 417 461 FIBRONECTIN TYPE-I.
 CC FT DOMAIN 462 504 FIBRONECTIN TYPE-I.
 CC FT DOMAIN 506 >522 FIBRONECTIN TYPE-I.
 CC FT SITE 304 306 CELL ATTACHMENT SITE (POTENTIAL).
 CC FT SITE 419 448 BY SIMILARITY.
 CC FT DISULFID 446 458 BY SIMILARITY.
 CC FT DISULFID 464 491 BY SIMILARITY.
 CC FT DISULFID 489 501 BY SIMILARITY.
 CC FT DISULFID 508 >522 BY SIMILARITY.
 CC FT CARBOHYD 321 321 POTENTIAL.
 CC FT VARSPLIC 205 205 D->E (IN ISOFORM (V+C)-).
 CC FT VARSPLIC 206 462 MISSING (IN ISOFORM (V+C)-).
 CC FT NON_TER 522 522
 CC SQ SEQUENCE 522 AA; 57700 MW; DB1D9A54C2BDAE26 CRC64;

Query Match 34.4%; Score 50.5; DB 1; Length 522;
 Best Local Similarity 42.9%; Pred. No. 5.8;
 Matches 9; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 2 pdvqdcpectlqenpfisqpg 22

Db 221 PEILDVPS-TVQKTPFITNG 240

RESULT 28

FINC_HORSE

ID FINC_HORSE STANDARD; PRT; 522 AA.

AC Q28377; Q28378;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE FIBRONECTIN (FN) (FRAGMENT).
 GN FNL.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Perissodactyla; Equidae; Equus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE; 96324983.
 RA Maclod J.N., Burton-Wurster N., Gu D.N., Lust G.;
 RT "Fibronectin mRNA splice variant in articular cartilage lacks bases
 encoding the V, III-15, and I-10 protein segments."; J.
 RL J. Biol. Chem. 271:18954-18960(1996).
 CC -!- FUNCTION: FIBRONECTINS BIND CELL SURFACES AND VARIOUS COMPOUNDS
 CC INCLUDING COLLAGEN, FIBRIN, HEPARIN, DNA, AND ACTIN. FIBRONECTINS
 CC ARE INVOLVED IN CELL ADHESION, CELL MOTILITY, OPSONIZATION, WOUND
 CC HEALING, AND MAINTENANCE OF CELL SHAPE.
 CC -!- FUNCTION: VARIANT (V+C)- IS PROBABLY INVOLVED IN MATRIX
 CC ORGANIZATION OF CARTILAGE.
 CC -!- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED
 CC VARIANTS, CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS;
 CC TO A LESSER EXTEND HOMODIMERS.
 CC -!- ALTERNATIVE PRODUCTS: EACH OF THE "EXTRA DOMAIN" & THE CONNECTING
 CC STRAND 3 ARE PRESENT IN SOME FORMS OF FIBRONECTIN AND ABSENT IN
 CC OTHERS. THESE DIFFERENCES ARE DUE TO ALTERNATIVE SPLICING. VARIANT
 CC (V+C)- LACKS REPEAT 15 OF FIBRONECTIN TYPE-III, REPEAT 10 OF
 CC FIBRONECTIN TYPE-I, AND THE CONNECTING STRAND 3.
 CC -!- TISSUE SPECIFICITY: PLASMA FN (SOLUBLE DIMERIC FORM) IS SECRETED
 CC BY HEPATOCYTES. CELLULAR FN (DIMERIC OR CROSS-LINKED MULTIMERIC
 CC FORMS), MADE BY FIBROBLASTS, EPITHELIAL AND OTHER CELL TYPES, IS
 CC DEPOSITED AS FIBRILS IN THE EXTRACELLULAR MATRIX. VARIANT (V+C)-
 CC IS THE MAJOR TRANSCRIPT IN ARTICULAR CARTILAGE; VERY LOW LEVELS IN
 CC LYMPH NODE, BONE, AORTA, AND SKIN; ABSENT FROM LIVER, SPLEEN,
 CC PLACENTA, CARDIAC MUSCLE, SKELETAL MUSCLE, STOMACH, SMALL
 CC INTESTINE, AND KIDNEY.
 CC -!- SIMILARITY: CONTAINS AT LEAST 4 FIBRONECTIN TYPE III DOMAINS.
 CC -!- SIMILARITY: CONTAINS AT LEAST 3 FIBRONECTIN TYPE I DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U52107; AAC48613.1; -;
 CC EMBL; U52108; AAC48614.1; -;
 CC HSP; P02751; 1FBR.
 CC DR PROSITE; PS01253; FIBRONECTIN_1; 2.
 CC DR PFAM; PF00039; fn1; 2.
 CC DR PFAM; PF00041; fn3; 3.
 CC KW Glycoprotein; Plasma; Heparin-binding; Acute phase; Cell adhesion;
 CC Repeat; Alternative splicing.
 CC FT NON_TER 1 1
 CC FT DOMAIN 1 204 HEPARIN-BINDING 2 (BY SIMILARITY).
 CC FT DOMAIN 325 >522 FIBRIN-BINDING 2 (BY SIMILARITY).
 CC FT DOMAIN 26 114 FIBRONECTIN TYPE-III.
 CC FT DOMAIN 115 204 FIBRONECTIN TYPE-III.
 CC FT DOMAIN 205 315 CONNECTING STRAND 3 (CS-3) (V REGION).
 CC FT DOMAIN 316 395 FIBRONECTIN TYPE-III.
 CC FT DOMAIN 417 461 FIBRONECTIN TYPE-I.
 CC FT DOMAIN 462 504 FIBRONECTIN TYPE-I.
 CC FT DOMAIN 506 >522 FIBRONECTIN TYPE-I.
 CC FT SITE 304 306 CELL ATTACHMENT SITE (POTENTIAL).
 CC FT SITE 419 448 BY SIMILARITY.
 CC FT DISULFID 446 458 BY SIMILARITY.
 CC FT DISULFID 464 491 BY SIMILARITY.
 CC FT DISULFID 489 501 BY SIMILARITY.
 CC FT DISULFID 508 >522 BY SIMILARITY.
 CC FT CARBOHYD 321 321 POTENTIAL.
 CC FT VARSPLIC 205 205 D->E (IN ISOFORM (V+C)-).
 CC FT VARSPLIC 206 462 MISSING (IN ISOFORM (V+C)-).
 CC FT NON_TER 522 522
 CC SQ SEQUENCE 522 AA; 57700 MW; DB1D9A54C2BDAE26 CRC64;


```

FT DISULFID 512 524
FT DISULFID 530 558
FT DISULFID 556 568
FT DISULFID 2085 2114
FT DISULFID 2112 2124
FT DISULFID 2130 2157
FT DISULFID 2155 2167
FT DISULFID 2174 2200
FT DISULFID 2198 2209
FT DISULFID 2246 2246
FT DISULFID 2250 2250
FT CARBOHYD 399 399
FT CARBOHYD 497 497
FT CARBOHYD 511 511
FT CARBOHYD 846 846
FT CARBOHYD 976 976
FT CARBOHYD 1213 1213
FT CARBOHYD 1987 1987
FT CARBOHYD 1943 1943
FT CARBOHYD 1944 1944
FT MOD_RES 2263 2263
SQ SEQUENCE 2265 AA; 249557 MW; C2D21D486F498D5C CRC64;

Query Match 34.4%; Score 50.5; DB 1; Length 2265;
Best Local Similarity 42.9%; Pred. No. 27;
Matches 9; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 2 pdvqdspectqlenpfqpg 22
   | | | | | | | | | |
DB 1887 PEILDVFS-TVQKTPFTTNG 1906

RESULT 30
FINC_MOUSE
ID FINC_MOUSE STANDARD; PRT: 2477 AA.
OC P11276; Q61568; Q61569; Q61567; Q64233;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE FIBRONECTIN PRECURSOR (FN) (FRAGMENTS).
GN FNI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RN SEQUENCE OF 1-28 FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 94131313.
RA Polly P., Nicholson R.C.;
RT "Sequence of the mouse fibronectin-encoding gene promoter region.";
RL Gene 137:353-354(1993).
[2]
RN SEQUENCE OF 562-834 FROM N.A.
RC STRAIN=NMRI;
RX MEDLINE; 95403556.
RA Talts J.F., Weller A., Ekblom M., Ekblom P.;
RT "Regulation of mesenchymal extracellular matrix protein synthesis by
transforming growth factor-beta and glucocorticoids in tumor
stroma.";
RL J. Cell Sci. 108:2153-2162(1995).
[3]
RN SEQUENCE OF 899-2376 FROM N.A.
RA Gorski G., Aros M., Norton P.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE OF 2375-2477 FROM N.A.
RX MEDLINE; 88124987.
RA Blatti S.P., Foster D.N., Ranganathan G., Moses H.L., Getz M.J.;
RT "Induction of fibronectin gene transcription and mRNA is a primary
response to growth-factor stimulation of AKR-2B cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:1119-1123(1988).
[5]

```

```

RP SEQUENCE OF 2375-2477 FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE; 93011702.
RA Khandjian E.W., Salomon C., Leonard N., Tremblay S., Turler H.;
RT "Fibronectin gene expression in proliferating, quiescent, and SV40-
infected mouse kidney cells.";
RL Exp. Cell Res. 202:464-470(1992).
[6]
RP STRUCTURE BY NMR OF 1447-1630.
RX MEDLINE; 98202578.
RA Copie V., Tomita Y., Akiyama S.K., Aota S., Yamada K.M., Venable R.M.,
RA Pastor R.W., Krueger S., Torchia D.A.;
RT "Solution structure and dynamics of linked cell attachment modules of
mouse fibronectin containing the RGD and synergy regions: comparison
with the human fibronectin crystal structure.";
RL J. Mol. Biol. 277:663-682(1998).
CC -!- FUNCTION: FIBRONECTINS BIND CELL SURFACES AND VARIOUS COMPOUNDS
INCLUDING COLLAGEN, FIBRIN, HEPARIN, DNA, AND ACTIN. FIBRONECTINS
ARE INVOLVED IN CELL ADHESION, CELL MOTILITY, OPSONIZATION, WOUND
HEALING, AND MAINTENANCE OF CELL SHAPE.
CC -!- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED
VARIANTS, CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS;
TO A LESSER EXTENT HOMODIMERS.
CC -!- ALTERNATIVE PRODUCTS: EACH OF THE "EXTRA DOMAIN" & THE CONNECTING
STRAND 3 ARE PRESENT IN SOME FORMS OF FIBRONECTIN AND ABSENT IN
OTHERS. THESE DIFFERENCES ARE DUE TO ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: PLASMA FN (SOLUBLE DIMERIC FORM) IS SECRETED
BY HEPATOCYTES. CELLULAR FN (DIMERIC OR CROSS-LINKED MULTIMERIC
FORMS), MADE BY FIBROBLASTS, EPITHELIAL AND OTHER CELL TYPES, IS
DEPOSITED AS FIBRILS IN THE EXTRACELLULAR MATRIX.
CC -!- SIMILARITY: CONTAINS 12 FIBRONECTIN TYPE I DOMAINS.
CC -!- SIMILARITY: CONTAINS 17 FIBRONECTIN TYPE II DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
EMBL; 227279; CAA80422.1; -
DR EMBL; X82402; CAA57796.1; -
DR EMBL; X93167; CAA63654.1; -
DR EMBL; M18194; AAA37636.1; -
DR EMBL; S45680; AAB23491.1; -
DR PIR; A31371; A31371.
DR PIR; C60597; C60597.
DR PDB; 1MFN; 29-APR-98.
DR PDB; 2MFN; 29-APR-98.
DR MGD; MGI-95566; FN1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; PARTIAL.
DR PROSITE; PS01253; FIBRONECTIN_1; 3.
DR PRAM; PF00039; fn1; 4.
DR PRAM; PF00041; fn3; 16.
KW Glycoprotein; Plasma; Heparin-binding; Acute phase; Phosphorylation;
KW Sulfatation; Cell adhesion; Repeat; Alternative splicing; Signal;
KW 3D-structure.
FT SIGNAL 1 32
FT CHAIN 33 2477
FT DOMAIN 53 273
FT DOMAIN 308 608
FT DNA_BIND 906 1171
FT DOMAIN 1357 1630
FT DOMAIN 1811 2081
FT DOMAIN 2296 2427
FT DOMAIN 51 96
FT DOMAIN 96 140
FT DOMAIN 140 185
FT DOMAIN 185 230
FT DOMAIN 230 272

```


RL Nucleic Acids Res. 12:5853-5868(1984).
RN [5]
RP SEQUENCE OF 1594-2386 FROM N.A.
RX MEDLINE; 85280409.
RA Bernard M.P., Kolbe M., Weil D., Chu M.-L.;
RT "Human cellular fibronectin: comparison of the carboxyl-terminal
RT portion with rat identifies primary structural domains separated by
RT hypervariable regions.";
RL Biochemistry 24:2698-2704(1985).
RN [6]
RP SEQUENCE OF 32-290.
RX MEDLINE; 84032463.
RA Garcia-Pardo A., Pearlstein E., Frangione B.;
RT "Primary structure of human plasma fibronectin. The 29,000-dalton
RT NH2-terminal domain.";
RL J. Biol. Chem. 258:12670-12674(1983).
RN [7]
RP SEQUENCE OF 309-608, AND COLLAGEN-BINDING.
RX MEDLINE; 87080265.
RA Owens R.J., Barallie F.E.;
RT "Mapping the collagen-binding site of human fibronectin by expression
RT in Escherichia coli.";
RL EMBO J. 5:2825-2830(1986).
RN [8]
RP SEQUENCE OF 1441-1548.
RX MEDLINE; 82265604.
RA Pierschbacher M.D., Ruoslahti E., Sundelin J., Lind P., Peterson P.A.;
RT "The cell attachment domain of fibronectin. Determination of the
RT primary structure.";
RL J. Biol. Chem. 257:9593-9597(1982).
RN [9]
RP SEQUENCE OF 1434-1537 FROM N.A.
RX MEDLINE; 83290929.
RA Oldberg A., Linney E., Ruoslahti E.;
RT "Molecular cloning and nucleotide sequence of a cDNA clone coding for
RT the cell attachment domain in human fibronectin.";
RL J. Biol. Chem. 258:10193-10196(1983).
RN [10]
RP SEQUENCE OF 1448-1540 FROM N.A.
RX MEDLINE; 86111901.
RA Oldberg A., Ruoslahti E.;
RT "Evolution of the fibronectin gene. Exon structure of cell attachment
RT domain.";
RL J. Biol. Chem. 261:2113-2116(1986).
RN [11]
RP SEQUENCE OF 1712-1739 FROM N.A.
RX MEDLINE; 87026578.
RA Sekiguchi K., Klos A.M., Kurachi K., Yoshitake S., Hakomori S.;
RT "Human liver fibronectin complementary DNAs: identification of two
RT different messenger RNAs possibly encoding the alpha and beta
RT subunits of plasma fibronectin.";
RL Biochemistry 25:4936-4941(1986).
RN [12]
RP CHARACTERIZATION OF FIBRIN-BINDING SITE 1.
RX MEDLINE; 95081153.
RA Rostagno A., Williams M.J., Baron M., Campbell I.D., Gold L.I.;
RT "Further characterization of the NH2-terminal fibrin-binding site on
RT fibronectin.";
RL J. Biol. Chem. 269:31938-31945(1994).
RN [13]
RP STRUCTURE BY NMR OF 32-92.
RX MEDLINE; 96069779.
RA Potts J.R., Phan I., Williams M.J., Campbell I.D.;
RT "High-resolution structural studies of the factor XIIIa crosslinking
RT site and the first type 1 module of fibronectin.";
RL Nat. Struct. Biol. 2:1946-1950(1995).
RN [14]
RP STRUCTURE BY NMR OF 182-275.
RX MEDLINE; 94141923.
RA Williams M.J., Phan I., Harvey T.S., Rostagno A., Gold L.I.,
RA Campbell I.D.;
RT "Solution structure of a pair of fibronectin type 1 modules with
RT fibrin binding activity.";
J. Mol. Biol. 235:1302-1311(1994).
RN [15]
RP STRUCTURE BY NMR OF 406-464.
RX MEDLINE; 98179558.
RA Sticht H., Pickford A.R., Potts J.R., Campbell I.D.;
RT "Solution structure of the glycosylated second type 2 module of
RT fibronectin.";
RL J. Mol. Biol. 276:177-187(1998).
RN [16]
RP STRUCTURE BY NMR OF 1447-1540.
RX MEDLINE; 93046665.
RA Main A.L., Harvey T.S., Baron M., Boyd J., Campbell I.D.;
RT "The three-dimensional structure of the tenth type III module of
RT fibronectin: an insight into RGD-mediated interactions.";
RL Cell 71:671-678(1992).
RN [17]
RP STRUCTURE BY NMR OF 1447-1540.
RX MEDLINE; 92162710.
RA Baron M., Main A.L., Driscoll P.C., Mardon H.J., Boyd J.,
RA Campbell I.D.;
RT "1H NMR assignment and secondary structure of the cell adhesion type
RT III module of fibronectin.";
RL Biochemistry 31:2068-2073(1992).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 1447-1535.
RX MEDLINE; 94166075.
RA Dickinson C.D., Veerapandian B., Dai X.-P., Hamlin R.C., Xuong N.-H.,
RA Ruoslahti E., Ely K.R.;
RT "Crystal structure of the tenth type III cell adhesion module of
RT human fibronectin.";
RL J. Mol. Biol. 236:1079-1092(1994).
RN [19]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 1721-1991.
RX MEDLINE; 99177162.
RA Sharma A., Askari J.A., Humphries M.J., Jones E.Y., Stuart D.I.;
RT "Crystal structure of a heparin- and integrin-binding segment of human
RT fibronectin.";
RL EMBO J. 18:1468-1479(1999).
RN [20]
RP FUNCTION: FIBRONECTINS BIND CELL SURFACES AND VARIOUS COMPOUNDS
CC INCLUDING COLLAGEN, FIBRIN, HEPARIN, DNA, AND ACTIN. FIBRONECTINS
CC ARE INVOLVED IN CELL ADHESION, CELL MOTILITY, OPSONIZATION, WOUND
CC HEALING, AND MAINTENANCE OF CELL SHAPE.
CC
CC -1- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED
CC VARIANTS, CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS;
CC TO A LESSER EXTENT HOMODIMERS.
CC
CC -1- ALTERNATIVE PRODUCTS: EACH OF THE "EXTRA DOMAIN" & THE CONNECTING
CC STRAND 3 ARE PRESENT IN SOME FORMS OF FIBRONECTIN AND ABSENT IN
CC OTHERS. THESE DIFFERENCES ARE DUE TO ALTERNATIVE SPLICING.
CC
CC -1- TISSUE SPECIFICITY: PLASMA FN (SOLUBLE DIMERIC FORM) IS SECRETED
CC BY HEPATOCYTES. CELLULAR FN (DIMERIC OR CROSS-LINKED MULTIMERIC
CC FORMS), MADE BY FIBROBLASTS, EPITHELIAL AND OTHER CELL TYPES, IS
CC DEPOSITED AS FIBRILS IN THE EXTRACELLULAR MATRIX.
CC
CC -1- SIMILARITY: CONTAINS 12 FIBRONECTIN TYPE I DOMAINS.
CC
CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE II DOMAINS.
CC
CC -1- SIMILARITY: CONTAINS 16 FIBRONECTIN TYPE III DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M15801; AAA53376.1; -;
CC EMBL; X02761; CAA26536.1; -;
CC EMBL; K00055; AAA52459.1; -;
CC EMBL; M10905; AAA52462.1; -;
CC EMBL; M12549; AAA58483.1; -;
CC EMBL; M14059; AAA52463.1; -;
CC PIR; A28460; FNHU.
CC PDB; 1TTT; 31-JAN-94.
CC PDB; 1TTG; 31-JAN-94.
CC


```

DR PDB: 1FNA: 30-APR-94.
DR PDB: 1FNE: 29-JAN-96.
DR PDB: 1FNR: 15-OCT-95.
DR PDB: 1FNH: 16-MAR-99.
DR PDB: 2FNE: PRELIMINARY.
DR MIN: 135600;
DR PRINTS: PR00012; FNTYPEI.
DR PRINTS: PR00013; FNTYPEII.
DR PRINTS: PR00014; FNTYPEIII.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS00023; FIBRONECTIN_2; 2.
DR PROSITE: PS01253; FIBRONECTIN_1; 12.
DR PFAM: PF00039; fn1; 12.
DR PFAM: PF00040; fn2; 2.
DR PFAM: PF00041; fn3; 16.
KW Glycoprotein; Plasma; Heparin-binding; Acute phase; Phosphorylation;
KW Sulfatation; Cell adhesion; Repeat; Alternative splicing; Signal;
KW 3D-structure.

Query Match 33.7%; Score 49.5; DB 1; Length 2386;
Best Local Similarity 42.9%; Pred. NO. 40;
Matches 9; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

Qy 2 pdvqdcpectqlqenpfsgp 22
Db 2008 PEILDVPS-TVQKIFPVTHPG 2027

RESULT 33
FNC_RAT
ID FNC_RAT STANDARD; PRT; 2477 AA.
AC P04937;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE FIBRONECTIN PRECURSOR (FN).
GN FNI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FISHER; TISSUE=LIVER;
RX MEDLINE; 88034951.
RA Schwarzbauer J.E., Patel R.S., Fonda D., Hynes R.O.;
RT "Multiple sites of alternative splicing of the rat fibronectin gene
transcript.";
RL EMO J. 6:2573-2580(1987).
RN [2]
RP SEQUENCE OF 1-139 AND 2382-2477 FROM N.A.
RC STRAIN=FISHER; TISSUE=LIVER;
RX MEDLINE; 88034950.
RA Patel R.S., Odermatt E., Schwarzbauer J.E., Hynes R.O.;
RT "Organization of the fibronectin gene provides evidence for exon
shuffling during evolution.";
RL EMO J. 6:2565-2572(1987).
RN [3]
RP SEQUENCE OF 1586-2477 FROM N.A.
RX MEDLINE; 84082067.
RA Schwarzbauer J.E., Tamkun J.W., Lemischka I.R., Hynes R.O.;
RT "Three different fibronectin mRNAs arise by alternative splicing
within the coding region.";
RL Cell 35:421-431(1983).
CC -!- FUNCTION: FIBRONECTINS BIND CELL SURFACES AND VARIOUS COMPOUNDS
INCLUDING COLLAGEN, FIBRIN, HEPARIN, DNA, AND ACTIN. FIBRONECTINS
ARE INVOLVED IN CELL ADHESION, CELL MOTILITY, OPSONIZATION, WOUND
HEALING, AND MAINTENANCE OF CELL SHAPE.
-!- SUBUNIT: MOSTLY HETERODIMERS OR MULTIMERS OF ALTERNATIVELY SPLICED
VARIANTS, CONNECTED BY 2 DISULFIDE BONDS NEAR THE CARBOXYL ENDS;
TO A LESSER EXTENT HOMODIMERS.
-!- ALTERNATIVE PRODUCTS: EACH OF THE "EXTRA DOMAIN" & THE CONNECTING
STRAND 3 ARE PRESENT IN SOME FORMS OF FIBRONECTIN AND ABSENT IN
OTHERS. THESE DIFFERENCES ARE DUE TO ALTERNATIVE SPLICING.

-!- TISSUE SPECIFICITY: PLASMA FN (SOLUBLE DIMERIC FORM) IS SECRETED
BY HEPATOCYTES. CELLULAR FN (DIMERIC OR CROSS-LINKED MULTIMERIC
FORMS), MADE BY FIBROBLASTS, EPITHELIAL AND OTHER CELL TYPES, IS
DEPOSITED AS FIBRILS IN THE EXTRACELLULAR MATRIX.
-!- SIMILARITY: CONTAINS 12 FIBRONECTIN TYPE I DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE II DOMAINS.
CC -!- SIMILARITY: CONTAINS 17 FIBRONECTIN TYPE III DOMAINS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
CC EMBL; X15906; CAA34020.1; -
DR EMBL; L29191; AAA41166.1; -
DR EMBL; L29191; AAA41166.1; JOINED.
DR EMBL; L29191; AAA41167.1; -
DR EMBL; L29191; AAA41167.1; JOINED.
DR EMBL; L29191; AAA41168.1; -
DR EMBL; L29191; AAA41168.1; JOINED.
DR EMBL; X05831; CAA29278.1; -
DR EMBL; X05832; CAA29279.1; -
DR EMBL; X05833; CAA29280.1; -
DR EMBL; X05834; CAA29281.1; -
DR PIR; A27252; A27252.
DR PIR; S00459; S00459.
DR PIR; S14428; S14428.
DR HSP; P02751; 1FBR.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00014; FNTYPEIII.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS01253; FIBRONECTIN_1; 12.
DR PFAM; PF00039; fn1; 12.
DR PFAM; PF00040; fn2; 2.
DR PFAM; PF00041; fn3; 17.
KW Glycoprotein; Plasma; Heparin-binding; Acute phase; Phosphorylation;
KW Sulfatation; Cell adhesion; Repeat; Alternative splicing; Signal.
FT SIGNAL 1 32
FT CHAIN 33 2477 FIBRONECTIN.
FT DOMAIN 53 273 FIBRIN- AND HEPARIN-BINDING 1.
FT DOMAIN 308 608 COLLAGEN-BINDING.
FT DNA_BIND 906 1171
FT DOMAIN 1357 1630 CELL-ATTACHMENT.
FT DOMAIN 1811 2081 HEPARIN-BINDING 2.
FT DOMAIN 2296 2427 FIBRIN-BINDING 2.
FT DOMAIN 51 91 FIBRONECTIN TYPE-I 1.
FT DOMAIN 96 139 FIBRONECTIN TYPE-I 2.
FT DOMAIN 140 183 FIBRONECTIN TYPE-I 3.
FT DOMAIN 185 229 FIBRONECTIN TYPE-I 4.
FT DOMAIN 230 274 FIBRONECTIN TYPE-I 5.
FT DOMAIN 306 345 FIBRONECTIN TYPE-I 6.
FT DOMAIN 345 404 FIBRONECTIN TYPE-II 1.
FT DOMAIN 405 459 FIBRONECTIN TYPE-II 2.
FT DOMAIN 468 511 FIBRONECTIN TYPE-I 7.
FT DOMAIN 516 558 FIBRONECTIN TYPE-I 8.
FT DOMAIN 559 602 FIBRONECTIN TYPE-I 9.
FT DOMAIN 609 706 FIBRONECTIN TYPE-III 1.
FT DOMAIN 707 808 FIBRONECTIN TYPE-III 2.
FT DOMAIN 809 903 FIBRONECTIN TYPE-III 3.
FT DOMAIN 904 984 FIBRONECTIN TYPE-III 4.
FT DOMAIN 995 1084 FIBRONECTIN TYPE-III 5.
FT DOMAIN 1085 1172 FIBRONECTIN TYPE-III 6.
FT DOMAIN 1173 1264 FIBRONECTIN TYPE-III 7.
FT DOMAIN 1265 1355 FIBRONECTIN TYPE-III 8.
FT DOMAIN 1356 1446 FIBRONECTIN TYPE-III 9.
FT DOMAIN 1447 1536 FIBRONECTIN TYPE-III 10.
FT DOMAIN 1537 1630 FIBRONECTIN TYPE-III 11.
FT DOMAIN 1631 1720 FIBRONECTIN TYPE-III 12.

```



```

FT DOMAIN 1721 1810 FIBRONECTIN TYPE-III 13 (EXTRA DOMAIN 2).
FT DOMAIN 1811 1902 FIBRONECTIN TYPE-III 14.
FT DOMAIN 1903 1991 FIBRONECTIN TYPE-III 15.
FT DOMAIN 1992 2081 FIBRONECTIN TYPE-III 16.
FT DOMAIN 2082 2201 CONNECTING STRAND 3 (CS-3) (V REGION).
FT DOMAIN 2202 2283 FIBRONECTIN TYPE-III 17.
FT DOMAIN 2284 2338 FIBRONECTIN TYPE-I 10.
FT DOMAIN 2339 2381 FIBRONECTIN TYPE-I 11.
FT DOMAIN 2382 2426 FIBRONECTIN TYPE-I 12.
FT SITE 1614 1616 CELL ATTACHMENT SITE.
FT SITE 2181 2183 CELL ATTACHMENT SITE.
FT DISULFID 53 79 BY SIMILARITY.
FT DISULFID 77 88 BY SIMILARITY.
FT DISULFID 98 126 BY SIMILARITY.
FT DISULFID 124 136 BY SIMILARITY.
FT DISULFID 142 170 BY SIMILARITY.
FT DISULFID 168 180 BY SIMILARITY.
FT DISULFID 187 216 BY SIMILARITY.
FT DISULFID 214 226 BY SIMILARITY.
FT DISULFID 232 261 BY SIMILARITY.
FT DISULFID 259 271 BY SIMILARITY.
FT DISULFID 308 335 BY SIMILARITY.
FT DISULFID 333 342 BY SIMILARITY.
FT DISULFID 360 386 BY SIMILARITY.
FT DISULFID 374 401 BY SIMILARITY.
FT DISULFID 420 446 BY SIMILARITY.
FT DISULFID 434 461 BY SIMILARITY.
FT DISULFID 470 498 BY SIMILARITY.
FT DISULFID 496 508 BY SIMILARITY.
FT DISULFID 518 545 BY SIMILARITY.
FT DISULFID 543 555 BY SIMILARITY.
FT DISULFID 561 589 BY SIMILARITY.
FT DISULFID 587 599 BY SIMILARITY.
FT DISULFID 2296 2325 BY SIMILARITY.
FT DISULFID 2323 2335 BY SIMILARITY.
FT DISULFID 2341 2368 BY SIMILARITY.
FT DISULFID 2366 2378 BY SIMILARITY.
FT DISULFID 2385 2409 BY SIMILARITY.
FT DISULFID 2407 2423 BY SIMILARITY.
FT DISULFID 2458 2458 INTERCHAIN (WITH 2462 OF OTHER CHAIN).
FT DISULFID 2462 2462 INTERCHAIN (WITH 2458 OF OTHER CHAIN).
FT CARBOHYD 430 430 POTENTIAL.
FT CARBOHYD 528 528 POTENTIAL.
FT CARBOHYD 542 542 POTENTIAL.
FT CARBOHYD 876 876 POTENTIAL.
FT CARBOHYD 1006 1006 POTENTIAL.
FT CARBOHYD 1243 1243 POTENTIAL.
FT CARBOHYD 1290 1290 POTENTIAL.
FT CARBOHYD 2198 2198 POTENTIAL.
FT MOD_RES 2475 PHOSPHORYLATION (BY SIMILARITY).
FT VARSPLIC 1720 1809 MISSING (IN ISOFORM FNIII-13 LESS).
FT VARSPLIC 2082 2106 MISSING (IN ISOFORM LAMBDA-RLF4-5).
FT VARSPLIC 2082 2200 MISSING (IN ISOFORM LAMBDA-RLF6).
FT CONFLICT 2318 2318 G -> A (IN REF. 3).
SQ SEQUENCE 2477 AA; 272510 MW; B4391A472ECEDEB5 CRC64;

Query Match 33.7%; Score 49.5; DB 1; Length 2477;
Best Local Similarity 42.9%; Pred. No. 41;
Matches 9; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 2 pdvqdcpectlqenffsqpg 22
|::| |::| |::| |
Db 2098 PEILDVPS-TVQKTFVTNPG 2117

RESULT 34
HUPJ_RHOCA
ID HUPJ_RHOCA STANDARD; PRT; 278 AA.
AC Q03009;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)

DE PROBABLE RUBREDOXIN HUPJ.
GN HUPJ.
OS Rhodobacter capsulatus (Rhodopseudomonas capsulata).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group; Rhodobacter.
OC [1]
RN SEQUENCE FROM N.A.
RP STRAIN=B10;
RX MEDLINE; 93268090.
RA Colbeau A., Richaud P., Toussaint B., Caballero F.J., Elster C., Delphin C., Smith R.L., Chabert J., Vignais P.M.;
RT "Organization of the genes necessary for hydrogenase expression in Rhodobacter capsulatus. Sequence analysis and identification of two hyp regulatory mutants.";
RL Mol. Microbiol. 8:15-29(1993).
RN [2]
RP PRELIMINARY SEQUENCE OF 172-278 FROM N.A.
RX MEDLINE; 91177833.
RA Xu H.W., Wall J.D.;
RT "Clustering of genes necessary for hydrogen oxidation in Rhodobacter capsulatus.";
RL J. Bacteriol. 173:2401-2405(1991).
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RUBREDOXIN FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HUPJ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z15089; CAA78802.1; .
CC DR EMBL; M55089; AAA72923.1; ALT_SEQ.
CC DR PIR; S25690; S25690.
CC DR HSSP; P00289; 7RXN.
CC DR PRINTS; PRO0163; RUBREDOXIN.
CC DR PROSITE; PS00202; RUBREDOXIN; 1.
CC DR PFAM; PF00301; rubredoxin; 1.
CC KW Electron transport; Iron; Metal-binding.
CC FT DOMAIN 1 87 RUBREDOXIN-LIKE.
CC FT METAL 28 28 IRON (BY SIMILARITY).
CC FT METAL 31 31 IRON (BY SIMILARITY).
CC FT METAL 61 61 IRON (BY SIMILARITY).
CC FT METAL 64 64 IRON (BY SIMILARITY).
CC SQ SEQUENCE 278 AA; 30329 MW; 4D750CB411393A8D CRC64;

Query Match 33.3%; Score 49; DB 1; Length 278;
Best Local Similarity 40.7%; Pred. No. 5;
Matches 11; Conservative 2; Mismatches 12; Indels 2; Gaps 1;

QY 2 pdvqdcpectlqenpff--sqgqapil 26
|::| |::| |::| |
Db 56 PEDWHCPNDAPKAQFIVSDPGAPAL 82

RESULT 35
GLHA_MORSA
ID GLHA_MORSA STANDARD; PRT; 117 AA.
AC Q91119;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GLYCOPROTEIN HORMONES ALPHA CHAIN PRECURSOR (GONADOTROPIN ALPHA CHAIN) (GTH-ALPHA).
OS Morone saxatilis (Striped bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha; Perciformes; Percolidae; Moronidae; Morone.
RN [1]
```


CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES ALPHA CHAIN
 CC FAMILY.
 CC HSP; P01215; IHRP.
 DR PRINTS; P00274; GLYCOHORMONE.
 DR PRINTS; P00438; GFCYSKNOT.
 DR PROSITE; P000779; GLYCOHORMONE_ALPHA_1; 1.
 DR PROSITE; P000780; GLYCOHORMONE_ALPHA_2; 1.
 DR PFAM; P00236; hormone6; 1.
 KW Hormone; Glycoprotein.
 FT DISULFID 11 34 BY SIMILARITY.
 FT DISULFID 14 63 BY SIMILARITY.
 FT DISULFID 31 84 BY SIMILARITY.
 FT DISULFID 35 86 BY SIMILARITY.
 FT DISULFID 62 89 BY SIMILARITY.
 FT CARBOHYD 55 55
 FT CARBOHYD 80 80
 SQ SEQUENCE 94 AA; 10665 MW; 29CD7C06F5D4BD0E CRC64;

Query Match 32.0%; Score 47; DB 1; Length 94;
 Best Local Similarity 57.1%; Pred. No. 3.1;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 cpecltqenpfssq 20
 | |||||:| ||:
 Db 11 CEECTLKNNVFSR 24

RESULT 38
 RPB9_YEAST
 ID RPB9_YEAST STANDARD; PRT; 122 AA.
 AC P27999;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE DNA-DIRECTED RNA POLYMERASE II 14.2 KD POLYPEPTIDE (EC 2.7.7.6)
 DE (B12.6).
 GN RPB9 OR YGL070C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomycetes.
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 46-62.
 RX MEDLINE; 92011681.
 RA Woychik N.A., Lane W.S., Young R.A.;
 RT "Yeast RNA polymerase II subunit RPB9 is essential for growth at
 RT temperature extremes";
 RL J. Biol. Chem. 266:19053-19055(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RX MEDLINE; 97435481.
 RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
 RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
 RT chromosome VII.";
 RL Yeast 13:1077-1090(1997).
 CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 CC SUBSTRATES.
 CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +
 CC RNA(N).
 CC -!- SUBUNIT: RNA POLYMERASE II CONSISTS OF 12 DIFFERENT SUBUNITS.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 CC -!- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE
 CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
 CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
 CC III FOR 5S AND TRNA GENES.
 CC -!- SIMILARITY: BELONGS TO THE ARCHAEABACTERIA RPB9 / EUKARYOTIC RPA12/
 CC RPB9 / RPB11 RNA POLYMERASE FAMILY.
 CC -!- SIMILARITY: BELONGS TO THE TRANSCRIPTION FACTOR S-II FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M73060; RAA34997.1; -;
 DR EMBL; Z72592; CAA96774.1; -;
 DR PIR; A41016; RNB929.
 DR SGD; L0001683; RPB9.
 DR PROSITE; PS00466; TFIIS; 1.
 DR PROSITE; PS01030; RNA_POL_15KD; 1.
 KW Transférase; DNA-directed RNA polymerase; Transcription;
 KW Nuclear protein; Zinc-finger.
 FT 2N_FING 7 32 C4-TYPE (POTENTIAL).
 FT 2N_FING 75 106 ZN-RIBBON (POTENTIAL).
 SQ SEQUENCE 122 AA; 14288 MW; A80D69678A722881 CRC64;

Query Match 32.0%; Score 47; DB 1; Length 122;
 Best Local Similarity 50.0%; Pred. No. 4.1;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 qdceptlqenpf 18
 ::||:| ||:
 Db 73 RECPKCHSRENVFF 86

RESULT 39
 HXC9_MOUSE
 ID HXC9_MOUSE STANDARD; PRT; 260 AA.
 AC P05633;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE HOMEOBOX PROTEIN HOX-C9 (HOX-3.2).
 GN HOXC9 OR HOXC-9 OR HOX-3.2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 92155117.
 RA Erselius J.R., Goulding M.D., Gruss P.;
 RT "Structure and expression pattern of the murine Hox-3.2 gene";
 RL Development 110:629-642(1990).
 RN [2]
 RP SEQUENCE OF 187-260 FROM N.A.
 RX MEDLINE; 88312579.
 RA Breier G., Dressler G.R., Gruss P.;
 RT "Primary structure and developmental expression pattern of Hox 3.1, a
 RT member of the murine Hox 3 homeobox gene cluster";
 RL EMBO J. 7:1329-1336(1988).
 RN [3]
 RP SEQUENCE OF 187-260 FROM N.A.
 RC STRAIN-CD-1; TISSUE-CONNECTIVE TISSUE;
 RX MEDLINE; 90349629.
 RA Angulewitsch A., Bieberich C., Bogarad L., Shashikant C., Ruddle F.H.;
 RT "Structural analysis of the Hox-3.1 transcription unit and the
 RT Hox-3.2-Hox-3.1 intergenic region";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:6428-6432(1990).
 CC -!- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
 CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
 CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 CC -!- SIMILARITY: BELONGS TO THE HOX-B FAMILY OF HOMEOBOX PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCINE DEHYDROGENASE [DECARBOXYLATING], MITOCHONDRIAL PRECURSOR
DE (EC 1.4.4.2) (GLYCINE DECARBOXYLASE) (GLYCINE CLEAVAGE SYSTEM P-
DE PROTEIN).
GN GDCSP.
OS Flaveria anomala.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Asteridae; euasterids II; Asterales; Asteraceae;
OC Asteroideae; Helanieae; Flaveria.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LEAF;
RA Nan Q., Bauwe H.;
RT "The GDCSP gene encoding P-protein of the glycine cleavage system in
RT the C3-C4 intermediate plant Flaveria anomala.";
RL (In) Plant Gene Register PGR98-004.
CC -|- FUNCTION: THE GLYCINE CLEAVAGE SYSTEM CATALYSES THE DEGRADATION OF
CC GLYCINE. THE P PROTEIN BINDS THE ALPHA-AMINO GROUP OF GLYCINE
CC THROUGH ITS PYRIDOXAL PHOSPHATE COFACTOR; CO(2) IS RELEASED AND
CC THE REMAINING METHYLAMINE MOIETY IS THEN TRANSFERRED TO THE
CC LIPOAMIDE COFACTOR OF THE H PROTEIN.
CC -|- CATALYTIC ACTIVITY: GLYCINE + LIPOYLPROTEIN = S-AMINOMETHYL-
CC DIHYDROLIPOYLPROTEIN + CO(2).
CC -|- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -|- SUBUNIT: HOMODIMER (BY SIMILARITY). THE GLYCINE CLEAVAGE SYSTEM IS
CC COMPOSED OF FOUR PROTEINS: P, T, L, AND H.
CC -|- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -|- SIMILARITY: TO OTHER GLYCINE CLEAVAGE SYSTEM P-PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z99762; CAB16911.1; -
CC OXIDOREDUCTASE; Mitochondrion; Pyridoxal phosphate; Transit peptide.
CC TRANSIT 1 63 MITOCHONDRION (POTENTIAL).
FT CHAIN 64 1034 GLYCINE DEHYDROGENASE [DECARBOXYLATING].
FT BINDING 770 770 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 1034 AA; 112695 MW; C2F424C76EF104BA CRC64;

Query Match 31.08; Score 45.5; DB 1; Length 1034;
Best Local Similarity 40.98; Pred. No. 62;
Matches 9; Conservative 5; Mismatches 7; Indels 1; Gaps 1;
QY 1 apdvqdc-pectlqenpfssqp 21
||:||||| :||:| :|
Db 518 APEVQDAIPSGLVRETPYLTHP 539

Query Match 31.08; Score 45.5; DB 1; Length 1034;
Best Local Similarity 40.98; Pred. No. 62;
Matches 9; Conservative 5; Mismatches 7; Indels 1; Gaps 1;
QY 1 apdvqdc-pectlqenpfssqp 21
||:||||| :||:| :|
Db 518 APEVQDAIPSGLVRETPYLTHP 539

RESULT 43
GDCSP_FLATR STANDARD; PRT; 1034 AA.
AC Q49852;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCINE DEHYDROGENASE [DECARBOXYLATING], MITOCHONDRIAL PRECURSOR
DE (EC 1.4.4.2) (GLYCINE DECARBOXYLASE) (GLYCINE CLEAVAGE SYSTEM P-
DE PROTEIN).
GN GDCSPA OR GDCSP.
OS Flaveria trinervia.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Asteridae; euasterids II; Asterales; Asteraceae;
OC Asteroideae; Helanieae; Flaveria.

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LEAF;
RA Cossu R., Bauwe H.;
RT "Two genes of the GDCSP gene family from the C4 plant Flaveria
RT trinervia: GDCSPA encoding P-protein and GDCSPB, a pseudogene.";
RL (In) Plant Gene Register PGR98-002.
CC -|- FUNCTION: THE GLYCINE CLEAVAGE SYSTEM CATALYSES THE DEGRADATION OF
CC GLYCINE. THE P PROTEIN BINDS THE ALPHA-AMINO GROUP OF GLYCINE
CC THROUGH ITS PYRIDOXAL PHOSPHATE COFACTOR; CO(2) IS RELEASED AND
CC THE REMAINING METHYLAMINE MOIETY IS THEN TRANSFERRED TO THE
CC LIPOAMIDE COFACTOR OF THE H PROTEIN.
CC -|- CATALYTIC ACTIVITY: GLYCINE + LIPOYLPROTEIN = S-AMINOMETHYL-
CC DIHYDROLIPOYLPROTEIN + CO(2).
CC -|- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -|- SUBUNIT: HOMODIMER (BY SIMILARITY). THE GLYCINE CLEAVAGE SYSTEM IS
CC COMPOSED OF FOUR PROTEINS: P, T, L, AND H.
CC -|- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -|- SIMILARITY: TO OTHER GLYCINE CLEAVAGE SYSTEM P-PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z99767; CAB16916.1; -
CC OXIDOREDUCTASE; Mitochondrion; Pyridoxal phosphate; Transit peptide.
CC TRANSIT 1 63 MITOCHONDRION (POTENTIAL).
FT CHAIN 64 1034 GLYCINE DEHYDROGENASE [DECARBOXYLATING].
FT BINDING 770 770 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 1034 AA; 112735 MW; 14A71076C05A5062 CRC64;

Query Match 31.08; Score 45.5; DB 1; Length 1034;
Best Local Similarity 40.98; Pred. No. 62;
Matches 9; Conservative 5; Mismatches 7; Indels 1; Gaps 1;
QY 1 apdvqdc-pectlqenpfssqp 21
||:||||| :||:| :|
Db 518 APEVQDAIPSGLVRETPYLTHP 539

RESULT 44
GCSA_FLAPR STANDARD; PRT; 1037 AA.
AC P49361;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLYCINE DEHYDROGENASE [DECARBOXYLATING] A, MITOCHONDRIAL PRECURSOR
DE (EC 1.4.4.2) (GLYCINE DECARBOXYLASE A) (GLYCINE CLEAVAGE SYSTEM P-
DE PROTEIN A).
GN GDCSPA.
OS Flaveria pringlei.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Asteridae; euasterids II; Asterales; Asteraceae;
OC Asteroideae; Helanieae; Flaveria.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LEAF;
RX MEDLINE; 95241630.
RA Bauwe H., Kopriwa S.;
RT "The gdcspa gene from Flaveria pringlei (Asteraceae).";
RL Plant Physiol. 107:655-655(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-LEAF;
RX MEDLINE; 94218395.
RA Kopriwa S., Bauwe H.;

"P-protein of glycine decarboxylase from Flaveria pringlei.";
Plant Physiol. 104:1077-1078(1994).
-!- FUNCTION: THE GLYCINE CLEAVAGE SYSTEM CATALYSES THE DEGRADATION OF GLYCINE. THE P PROTEIN BINDS THE ALPHA-AMINO GROUP OF GLYCINE THROUGH ITS PYRIDOXAL PHOSPHATE COFACTOR; CO(2) IS RELEASED AND THE REMAINING METHYLAMINE MOIETY IS THEN TRANSFERRED TO THE LIPOAMIDE COFACTOR OF THE H PROTEIN.
-!- CATALYTIC ACTIVITY: GLYCINE + LIPOYLPROTEIN = S-AMINOMETHYL-DIHYDROLIPOYLPROTEIN + CO(2).
-!- COFACTOR: PYRIDOXAL PHOSPHATE.
-!- SUBUNIT: HOMODIMER (BY SIMILARITY). THE GLYCINE CLEAVAGE SYSTEM IS COMPOSED OF FOUR PROTEINS: P, T, L, AND H.
-!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
-!- SIMILARITY: TO OTHER GLYCINE CLEAVAGE SYSTEM P-PROTEIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

DR EMBL; Z36879; CAA85353.1; -;
KW Multigene family.
FT TRANSIT 1 66 MITOCHONDRION (POTENTIAL).
FT CHAIN 67 1037 GLYCINE DEHYDROGENASE [DECARBOXYLATING] A.
FT BINDING 773 773 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT CONFLICT 2 2 E -> D (IN REF. 2).
FT CONFLICT 495 495 T -> I (IN REF. 2).
SQ SEQUENCE 1037 AA; 113031 MW; ED248FA227F9E0F3 CRC64;

Query Match 31.0%; Score 45.5; DB 1; Length 1037;
Best Local Similarity 40.9%; Pred. No. 62;
Matches 9; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

QY 1 apdvqdc-pectlqenpfssqp 21
II:III I :I I I I
Db 521 APEVQDAIPSGLVRETPYLTTP 542

RESULT 45
VE5_RHPV1
ID VE5_RHPV1 STANDARD; PRT; 157 AA.
AC P24834;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PROBABLE E5 PROTEIN.
GN E5.
OS Rhesus papillomavirus type 1 (Rhpv 1).
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Papillomavirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91135018.
RA Ostrow R.S., Labresh K.V., Faras A.J.;
RT "Characterization of the complete Rhpv 1 genomic sequence and an integration locus from a metastatic tumor.";
RL Virology 181:424-429(1991).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

DR EMBL; M60184; AAA79315.1; ALT_SEQ.

DR EMBL; M60184; AAA79316.1; ALT_SEQ.
KW PIR; F38503; W5WLR1.
SQ Early protein.
SEQUENCE 157 AA; 17398 MW; AC7AA67158844686 CRC64;

Query Match 30.6%; Score 45; DB 1; Length 157;
Best Local Similarity 53.8%; Pred. No. 10;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 dvqdcpectlqen 15
I :I I I I I I I
Db 104 DTPACPQCGLQON 116

RESULT 46
PUR2_ARATH
ID PUR2_ARATH STANDARD; PRT; 529 AA.
AC P52420;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PHOSPHORIBOSYLAMINE--GLYCINE LIGASE PRECURSOR (EC 6.3.4.13) (GARS)
DE (GLYCINAMIDE RIBONUCLEOTIDE SYNTHETASE) (PHOSPHORIBOSYLGLYCINAMIDE SYNTHETASE) (FRAGMENT).
GN PUR2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
OC Arabidopsis.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. COLUMBIA;
RX MEDLINE; 95004650.
RA Schnorr K.M., Nygaard P., Laloue M.;
RT "Molecular characterization of Arabidopsis thaliana cDNAs encoding three purine biosynthetic enzymes.";
RL Plant J. 6:113-121(1994).
CC -!- CATALYTIC ACTIVITY: ATP + 5-PHOSPHORIBOSYLAMINE + GLYCINE = ADP + ORTHOPHOSPHATE + 5'-PHOSPHORIBOSYLGLYCINAMIDE.
CC -!- PATHWAY: SECOND STEP IN DE NOVO PURINE BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- SIMILARITY: TO OTHER GARS FROM BACTERIA AND EUKARYOTES.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

DR EMBL; X74766; CAA52778.1; -;
DR PROSITE; PS00184; GARS; 1.
DR PFW; PF01071; GARS; 1.
KW Purine biosynthesis; Ligase; Chloroplast; Transit peptide.
FT NON_TER 1 1
FT TRANSIT <1 ?
FT CHAIN ? 529 CHLOROPLAST (POTENTIAL).
FT SEQUENCE 529 AA; 56358 MW; 7BEF19AD065D3B76 CRC64;

Query Match 30.6%; Score 45; DB 1; Length 529;
Best Local Similarity 52.9%; Pred. No. 37;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 8 pectlqenpfssqp 24
I :I I I I I I I
Db 40 PTTTQLNPFPSDPS 56

RESULT 47


```

CCD2 CAEEL
ID CCD2 CAEEL STANDARD; PRT; 360 AA.
AC P35799; Q22477;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CUTICLE COLLAGEN DPY-2 PRECURSOR.
GN DPY-2 OR T14B4.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE; 94060446.
RA Levy A.D., Yang J., Kramer J.M.;
RT "Molecular and genetic analyses of the Caenorhabditis elegans dpy-2
RT and dpy-10 collagen genes: a variety of molecular alterations affect
RT organismal morphology.";
RL Mol. Biol. Cell 4:803-817(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Favell T.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NEMATODE CUTICLES ARE COMPOSED LARGELY OF COLLAGEN-LIKE
CC PROTEINS. THE CUTICLE FUNCTIONS BOTH AS AN EXOSKELETON AND AS A
CC BARRIER TO PROTECT THE WORM FROM ITS ENVIRONMENT.
CC -!- SUBUNIT: COLLAGEN POLYPEPTIDE CHAINS ARE COMPLEXED WITHIN THE
CC CUTICLE BY DISULFIDE BONDS AND OTHER TYPES OF COVALENT CROSS-
CC LINKS.
CC -!- DISEASE: MUTATIONS IN DPY-2 AFFECTS THE BODY SHAPE.
CC -!- SIMILARITY: TO OTHER COLLAGENS. STRONG, TO OTHER CUTICLE
CC COLLAGENS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L12706; AAA17398.1; -.
CC EMBL; U50191; AAA91237.1; -.
CC WORMPEP; T14B4.6; CE04953.
CC PFAM; PF01391; Collagen; 2.
CC Cuticle; Connective tissue; Repeat; Multigene family; Collagen;
CC Signal.
CC SIGNAL 1 ?
CC CHAIN ? 360 POTENTIAL.
CC DOMAIN 123 ? CUTICLE COLLAGEN DPY-2.
CC DOMAIN 123 152 TRIPLE-HELICAL REGION.
CC DOMAIN 174 230 TRIPLE-HELICAL REGION.
CC DOMAIN 238 303 TRIPLE-HELICAL REGION.
CC VARIANT 129 129 G -> E (IN DPY2(SC38)).
CC VARIANT 183 183 G -> R (IN DPY2(E8)).
CC VARIANT 247 247 G -> R (IN DPY2(E489)).
CC VARIANT 253 253 G -> R (IN DPY2(Q292)).
CC CONFLICT 355 360 RRIRKW -> DGVGNESVSNHNKNGSYHLRFTQ (IN
CC REF. 2).
CC SEQUENCE 360 AA; 36738 MW; B827960838A22578 CRC64;

```

```

Query Match 30.3%; Score 44.5; DB 1; Length 360;
Best Local Similarity 29.0%; Pred. No. 29;
Matches 9; Conservative 6; Mismatches 9; Indels 7; Gaps 1;
QY 1 apdvqdcpectlq-----enpfsgqgap 24
Db 110 SPQFQECPCACIPGERGPSGSLPALPGAP 140

```

RESULT 48

```

GCST_BOVIN
ID GCST_BOVIN STANDARD; PRT; 397 AA.
AC P25285;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE AMINOMETHYLTRANSFERASE PRECURSOR (EC 2.1.2.10) (GLYCINE CLEAVAGE
DE SYSTEM T PROTEIN).
GN AMT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-55.
RC TISSUE=LIVER;
RX MEDLINE; 91161577.
RA Okamura-Ikeda K., Fujiwara K., Yamamoto M., Hiraga K., Motokawa Y.;
RT "Isolation and sequence determination of cDNA encoding T-protein of
RT the glycine cleavage system.";
RL J. Biol. Chem. 266:4917-4921(1991).
CC -!- FUNCTION: THE GLYCINE CLEAVAGE SYSTEM CATALYZES THE DEGRADATION OF
CC GLYCINE.
CC -!- CATALYTIC ACTIVITY: (6S)-TETRAHYDROFOLATE +
CC S-AMINOMETHYLDIHYDROLYPOYLPROTEIN - (5R)-5,10-
CC METHYLENETETRAHYDROFOLATE + NH(3) + DIHYDROLIPOYLPROTEIN.
CC -!- SUBUNIT: THE GLYCINE CLEAVAGE SYSTEM IS COMPOSED OF FOUR PROTEINS:
CC P, T, L, AND H.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M59799; AAA30786.1; -.
CC PIR; A23707; A23707.
CC PRAM; PF01571; GCV.T. 1.
CC TRANSFERASE; Aminotransferase; Mitochondrion; Transit peptide.
CC TRANSIT 1 22 MITOCHONDRION.
CC FT CHAIN 23 397 AMINOMETHYLTRANSFERASE.
CC SEQUENCE 397 AA; 42868 MW; 3FE5D598B78AB6FD CRC64;

```

```

Query Match 30.3%; Score 44.5; DB 1; Length 397;
Best Local Similarity 37.0%; Pred. No. 32;
Matches 10; Conservative 6; Mismatches 4; Indels 7; Gaps 2;

```

```

QY 7 cpecltqen-----pf-fsqgqapil 26
Db 345 CPSPCLAKKNVANGVPIEYSRPGTPLL 371

```

RESULT 49

```

FLGI_PSEPU
ID FLGI_PSEPU STANDARD; PRT; 367 AA.
AC Q52082;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FLAGELLAR P-RING PROTEIN PRECURSOR.
GN FLGI.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;
OC Pseudomonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PAW8;
RX MEDLINE; 95005455.
RA Winstanley C., Morgan J.A., Pickup R.W., Saunders J.R.;

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 14, 2000, 09:35:35 ; Search time 50.59 Seconds
(without alignments)

35.633 Million cell updates/sec

Title: ALPHA-CHAIN

Perfect score: 147

Sequence: 1 apdvqdeptclqenpfsgqapil 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225878

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 99%

Listing first 1000 summaries

Database :

SPTREMBL12:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mnc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	56.5	120	6 077752	O77752 trichosurus
2	82	55.8	120	6 09XSW8	O9XSW8 canis famil
3	77	52.4	120	6 046642	O46642 equus burch
4	73	49.7	107	13 090286	O90286 carassius a
5	73	49.7	108	13 090287	O90287 carassius a
6	73	49.7	116	13 09XGP3	O9XGP3 ictalurus p
7	73	49.7	119	13 091370	O91370 oncorhynch
8	55.5	37.8	189	11 097722	P97722 mus musculu
9	52	35.4	1523	11 088270	O88270 rattus norv
10	50.5	34.4	351	6 002816	O02816 oryctolagus
11	50.5	34.4	667	10 092P16	O92P16 arabidopsis
12	50	34.0	490	13 042492	O42492 fugu rubrip
13	49.5	33.7	212	4 095609	O95609 homo sapien
14	49.5	33.7	379	4 095617	O95617 homo sapien
15	47	32.0	55	1 058150	O58150 methanococc
16	47	32.0	114	5 09XW25	O9XW25 caenorhabdi
17	47	32.0	1113	5 027481	O27481 caenorhabdi
18	46.5	31.6	242	5 019919	O19919 caenorhabdi
19	46.5	31.6	746	2 066902	O66902 aquifex aeo

20	46	31.3	227	10	O22474	O22474 malus domes
21	45.5	31.0	433	4	O9Y315	O9Y315 homo sapien
22	45.5	31.0	3247	12	O65553	O65553 bovine herp
23	45	30.6	164	10	O23462	O23462 arabidopsis
24	45	30.6	190	12	O9WGB5	O9WGB5 herpesvirus
25	45	30.6	846	4	O9Y4K4	O9Y4K4 homo sapien
26	45	30.6	859	4	O9Y561	O9Y561 homo sapien
27	45	30.6	2004	5	O61222	O61222 caenorhabdi
28	44.5	30.3	228	4	O14873	O14873 homo sapien
29	44.5	30.3	318	4	O15162	O15162 homo sapien
30	44.5	30.3	326	13	O90771	O90771 gallus gall
31	44.5	30.3	531	2	O47944	O47944 gluconobact
32	44.5	30.3	572	5	O19594	O19594 caenorhabdi
33	44.5	30.3	724	5	O61225	O61225 sycon rapha
34	44.5	30.3	819	4	O12851	O12851 homo sapien
35	44.5	30.3	1013	4	O43897	O43897 homo sapien
36	44	29.9	55	2	O05894	O05894 mycobacteri
37	44	29.9	70	9	O92X75	O92X75 mycobacteri
38	44	29.9	393	1	O9Y9K2	O9Y9K2 aeropyrum p
39	44	29.9	601	5	O9Y128	O9Y128 drosophila
40	44	29.9	745	13	O91701	O91701 xenopus lae
41	44	29.9	753	4	O14830	O14830 homo sapien
42	44	29.9	821	11	O61161	O61161 mus musculu
43	44	29.9	2180	5	O01768	O01768 caenorhabdi
44	44	29.9	4545	11	O61291	O61291 mus musculu
45	43.5	29.6	735	13	O57381	O57381 xenopus lae
46	43.5	29.6	977	13	O91925	O91925 xenopus lae
47	43.5	29.6	1019	13	O57382	O57382 xenopus lae
48	43	29.3	272	2	O47891	O47891 fremyella d
49	43	29.3	307	11	O70233	O70233 mus musculu
50	43	29.3	421	5	O23852	O23852 dictyosteli
51	43	29.3	511	3	O59714	O59714 schizosacch
52	43	29.3	572	5	O20715	O20715 caenorhabdi
53	43	29.3	750	10	O81880	O81880 arabidopsis
54	43	29.3	812	4	O15793	O15793 homo sapien
55	43	29.3	812	5	O21623	O21623 caenorhabdi
56	43	29.3	825	4	O95644	O95644 homo sapien
57	43	29.3	877	5	O23853	O23853 dictyosteli
58	43	29.3	915	1	O58035	O58035 pyrococcus
59	43	29.3	1119	5	O21948	O21948 caenorhabdi
60	43	29.3	1266	10	O9XET3	O9XET3 lycopersico
61	42.5	28.9	254	5	O22644	O22644 caenorhabdi
62	42.5	28.9	570	10	O65471	O65471 arabidopsis
63	42.5	28.9	1022	13	O57460	O57460 brachydanio
64	42	28.6	100	2	O9X8E4	O9X8E4 streptomyce
65	42	28.6	114	13	O91371	O91371 oncorhynch
66	42	28.6	184	10	O22801	O22801 arabidopsis
67	42	28.6	204	11	O54730	O54730 mus musculu
68	42	28.6	234	11	O54731	O54731 mus musculu
69	42	28.6	242	11	O920Y0	O920Y0 mus musculu
70	42	28.6	289	12	O85066	O85066 peanut stun
71	42	28.6	332	2	O53175	O53175 rhodobacter
72	42	28.6	334	11	O921F6	O921F6 mus musculu
73	42	28.6	341	10	O92F06	O92F06 arabidopsis
74	42	28.6	346	11	O9WV28	O9WV28 mus musculu
75	42	28.6	361	4	O9Y611	O9Y611 homo sapien
76	42	28.6	364	12	O98288	O98288 molluscum c
77	42	28.6	369	10	O9XGF0	O9XGF0 triticum du
78	42	28.6	381	1	O27289	O27289 methanobact
79	42	28.6	381	10	O92N90	O92N90 triticum ae
80	42	28.6	384	1	O59044	O59044 methanococc
81	42	28.6	421	4	O9Y612	O9Y612 homo sapien
82	42	28.6	442	5	O04076	O04076 trypanosoma
83	42	28.6	479	2	O46877	O46877 escherichia
84	42	28.6	552	2	O50221	O50221 thiobacillu
85	42	28.6	601	5	O17552	O17552 caenorhabdi
86	42	28.6	609	3	O9Y746	O9Y746 kluyveromyc
87	42	28.6	612	5	O17206	O17206 caenorhabdi
88	42	28.6	684	5	P91132	P91132 caenorhabdi
89	42	28.6	771	4	O75147	O75147 homo sapien
90	42	28.6	839	5	P91583	P91583 ciona intes
91	42	28.6	1028	11	O64541	O64541 rattus norv
92	42	28.6	1086	4	O75808	O75808 homo sapien

93	42	28.6	1131	5	Q18529	Q18529 caenorhabdi	166	40	27.2	209	2	P94634	P94634 corynebacte
94	42	28.6	6048	5	Q23020	Q23020 caenorhabdi	167	40	27.2	216	13	Q9W7C7	Q9W7C7 oryzias lat
95	42	28.6	6831	5	Q23550	Q23550 caenorhabdi	168	40	27.2	275	13	Q93492	Q93492 cyprinus ca
96	42	28.6	7160	5	Q23551	Q23551 caenorhabdi	169	40	27.2	283	2	Q9W185	Q9W185 thermotoga
97	41.5	28.2	211	5	Q45522	Q45522 caenorhabdi	170	40	27.2	289	4	Q9Y5B1	Q9Y5B1 homo sapien
98	41.5	28.2	328	2	Q25855	Q25855 helicobacte	171	40	27.2	297	2	Q50009	Q50009 mycobacteri
99	41.5	28.2	378	6	Q9XS10	Q9XS10 bos indicus	172	40	27.2	297	2	Q52897	Q52897 rhodococcus
100	41.5	28.2	383	5	Q61710	Q61710 caenorhabdi	173	40	27.2	297	2	Q92N13	Q92N13 rhodococcus
101	41.5	28.2	386	9	Q80244	Q80244 mycoplasma	174	40	27.2	311	5	Q76712	Q76712 caenorhabdi
102	41.5	28.2	387	2	Q52861	Q52861 bacillus su	175	40	27.2	323	10	Q41547	Q41547 triticum ae
103	41.5	28.2	387	2	Q32062	Q32062 bacillus su	176	40	27.2	328	13	Q9YHX3	Q9YHX3 brachydanio
104	41.5	28.2	425	5	Q17400	Q17400 caenorhabdi	177	40	27.2	342	3	Q59681	Q59681 schizosacch
105	41.5	28.2	431	13	Q05467	Q05467 gallus gall	178	40	27.2	357	13	Q57317	Q57317 gallus gall
106	41.5	28.2	505	4	Q99520	Q99520 homo sapien	179	40	27.2	366	10	Q9XG59	Q9XG59 streptomyce
107	41.5	28.2	506	4	Q15346	Q15346 homo sapien	180	40	27.2	391	2	P72463	P72463 streptomyce
108	41.5	28.2	534	4	Q9Y3R6	Q9Y3R6 homo sapien	181	40	27.2	393	2	Q54877	Q54877 streptococc
109	41.5	28.2	600	5	Q17401	Q17401 caenorhabdi	182	40	27.2	407	1	Q26964	Q26964 methanobact
110	41.5	28.2	1012	11	Q9WVM6	Q9WVM6 mus musculu	183	40	27.2	418	13	Q90410	Q90410 brachydanio
111	41.5	28.2	1015	4	Q9Y617	Q9Y617 homo sapien	184	40	27.2	419	6	Q97925	Q97925 macropus gi
112	41.5	28.2	1606	13	Q57613	Q57613 gallus gall	185	40	27.2	421	10	Q64489	Q64489 arabidopsis
113	41	27.9	80	4	Q9Y2X5	Q9Y2X5 homo sapien	186	40	27.2	422	12	Q66733	Q66733 equine infe
114	41	27.9	138	4	Q16163	Q16163 homo sapien	187	40	27.2	423	5	Q9XUX5	Q9XUX5 caenorhabdi
115	41	27.9	209	8	Q9XNR7	Q9XNR7 pylaiella l	188	40	27.2	474	12	Q89542	Q89542 bovine herp
116	41	27.9	231	2	P71973	P71973 mycobacteri	189	40	27.2	478	2	Q84375	Q84375 chlamydia t
117	41	27.9	278	2	Q86062	Q86062 pseudomonas	190	40	27.2	491	10	Q22640	Q22640 raphanus sa
118	41	27.9	283	5	Q17458	Q17458 caenorhabdi	191	40	27.2	493	6	Q28617	Q28617 oryctolagus
119	41	27.9	294	5	Q17460	Q17460 caenorhabdi	192	40	27.2	494	4	Q43744	Q43744 homo sapien
120	41	27.9	334	11	Q70367	Q70367 rattus norv	193	40	27.2	540	1	Q27095	Q27095 methanobact
121	41	27.9	355	11	Q62295	Q62295 mus musculu	194	40	27.2	540	1	Q50576	Q50576 methanobact
122	41	27.9	370	2	Q9X8K5	Q9X8K5 streptomyce	195	40	27.2	576	11	Q61503	Q61503 mus musculu
123	41	27.9	453	10	Q49285	Q49285 arabidopsis	196	40	27.2	615	4	Q95754	Q95754 homo sapien
124	41	27.9	455	4	Q14250	Q14250 homo sapien	197	40	27.2	623	12	Q56659	Q56659 equine herp
125	41	27.9	457	11	Q92129	Q92129 rattus norv	198	40	27.2	685	5	Q44402	Q44402 caenorhabdi
126	41	27.9	474	3	Q94432	Q94432 schizosacch	199	40	27.2	692	5	Q77069	Q77069 drosophila
127	41	27.9	504	2	Q87783	Q87783 neisseria m	200	40	27.2	722	1	Q59644	Q59644 sulfolobus
128	41	27.9	683	8	Q9XP58	Q9XP58 triticum ae	201	40	27.2	745	5	Q61458	Q61458 drosophila
129	41	27.9	740	1	Q58078	Q58078 pyrococcus	202	40	27.2	756	3	Q93835	Q93835 neurospora
130	41	27.9	764	13	Q9W6G6	Q9W6G6 brachydanio	203	40	27.2	757	12	Q37364	Q37364 carrot red
131	41	27.9	771	5	Q76259	Q76259 drosophila	204	40	27.2	771	5	Q62609	Q62609 drosophila
132	41	27.9	792	6	Q19061	Q19061 saguinus oe	205	40	27.2	776	11	Q92143	Q92143 rattus norv
133	41	27.9	886	10	Q22016	Q22016 cylindrothe	206	40	27.2	777	11	Q92123	Q92123 mus musculu
134	41	27.9	946	10	Q22015	Q22015 cylindrothe	207	40	27.2	959	2	Q87815	Q87815 alcaligenes
135	41	27.9	1067	12	P89904	P89904 chimpanzee	208	40	27.2	1021	4	Q93033	Q93033 homo sapien
136	41	27.9	1271	5	Q21789	Q21789 caenorhabdi	209	40	27.2	1127	5	Q62205	Q62205 caenorhabdi
137	41	27.9	1294	4	Q14529	Q14529 homo sapien	210	40	27.2	1138	12	Q9W7U6	Q9W7U6 equine infe
138	41	27.9	1421	10	Q80907	Q80907 arabidopsis	211	40	27.2	1143	12	Q10696	Q10696 hog cholera
139	41	27.9	1502	11	Q88269	Q88269 rattus norv	212	40	27.2	1146	12	Q89468	Q89468 equine infe
140	41	27.9	1651	4	Q9Y6N7	Q9Y6N7 homo sapien	213	40	27.2	1146	12	Q89472	Q89472 equine infe
141	41	27.9	1879	5	Q18210	Q18210 caenorhabdi	214	40	27.2	1165	5	Q62206	Q62206 caenorhabdi
142	40.5	27.6	131	10	Q42350	Q42350 arabidopsis	215	40	27.2	1354	13	Q9W6B2	Q9W6B2 xenopus lae
143	40.5	27.6	287	4	Q76101	Q76101 homo sapien	216	40	27.2	1717	5	Q26566	Q26566 schistosoma
144	40.5	27.6	411	12	Q36409	Q36409 alcelaphine	217	40	27.2	3898	12	Q68534	Q68534 hog cholera
145	40.5	27.6	482	2	Q9X100	Q9X100 thermotoga	218	40	27.2	3898	12	Q96891	Q96891 hog cholera
146	40.5	27.6	538	4	Q15291	Q15291 homo sapien	219	40	27.2	3898	12	Q92364	Q92364 hog cholera
147	40.5	27.6	552	12	Q72909	Q72909 fowipox vir	220	40	27.2	4498	13	Q93291	Q93291 fugu rubrip
148	40.5	27.6	652	4	Q9Y660	Q9Y660 homo sapien	221	39.5	26.9	111	2	Q92522	Q92522 streptomyce
149	40.5	27.6	707	11	P97860	P97860 mus musculu	222	39.5	26.9	218	5	Q9XXF0	Q9XXF0 caenorhabdi
150	40.5	27.6	717	4	Q9Y2W0	Q9Y2W0 homo sapien	223	39.5	26.9	220	2	Q95989	Q95989 myxococcus
151	40.5	27.6	1013	11	Q62381	Q62381 mus musculu	224	39.5	26.9	231	2	Q9X1U0	Q9X1U0 thermotoga
152	40.5	27.6	1037	10	Q82642	Q82642 arabidopsis	225	39.5	26.9	311	5	Q93640	Q93640 caenorhabdi
153	40.5	27.6	1077	3	Q05549	Q05549 saccharomyc	226	39.5	26.9	329	12	Q9YVK0	Q9YVK0 melanoplus
154	40.5	27.6	1256	11	Q35158	Q35158 rattus norv	227	39.5	26.9	336	2	Q31619	Q31619 bacillus su
155	40.5	27.6	1323	11	Q62645	Q62645 rattus norv	228	39.5	26.9	360	5	Q9XUJ9	Q9XUJ9 caenorhabdi
156	40.5	27.6	1323	11	Q63381	Q63381 rattus norv	229	39.5	26.9	449	4	Q15113	Q15113 homo sapien
157	40.5	27.6	1323	11	Q63382	Q63382 rattus norv	230	39.5	26.9	449	4	Q14550	Q14550 homo sapien
158	40.5	27.6	1336	4	Q15399	Q15399 homo sapien	231	39.5	26.9	509	2	Q31240	Q31240 anabaena sp
159	40.5	27.6	1356	11	Q63729	Q63729 rattus norv	232	39.5	26.9	532	5	Q17500	Q17500 caenorhabdi
160	40.5	27.6	1406	4	Q15082	Q15082 homo sapien	233	39.5	26.9	798	4	Q9Y5E9	Q9Y5E9 homo sapien
161	40.5	27.6	1743	5	Q9XW55	Q9XW55 caenorhabdi	234	39.5	26.9	970	10	Q80790	Q80790 arabidopsis
162	40	27.2	77	5	Q77417	Q77417 anisakis si	235	39.5	26.9	1072	3	Q13807	Q13807 schizosacch
163	40	27.2	159	2	Q46801	Q46801 escherichia	236	39.5	26.9	1246	4	Q76046	Q76046 homo sapien
164	40	27.2	180	11	Q35445	Q35445 mus musculu	237	39.5	26.9	1250	11	Q88971	Q88971 mus musculu
165	40	27.2	193	2	Q92HD8	Q92HD8 buchnera ap	238	39.5	26.9	1833	11	Q08999	Q08999 mus musculu

239	39.5	26.9	1873	2	Q924N7	Q9z4n7 enterococc	312	39	26.5	1345	12	Q69063	Q69063 human herpe
240	39.5	26.9	1993	5	P90670	P90670 aplysia cal	313	39	26.5	1345	12	Q9WT16	Q9wt16 human herpe
241	39	26.5	55	11	Q9WUN4	Q9wun4 rattus norv	314	39	26.5	1421	5	Q05644	Q05644 plasmodium
242	39	26.5	80	5	Q21102	Q21102 caenorhabdi	315	39	26.5	1475	5	Q25842	Q25842 plasmodium
243	39	26.5	93	9	Q92X32	Q92x32 mycobacteri	316	39	26.5	1527	4	Q75621	Q75621 homo sapien
244	39	26.5	112	2	Q92FQ1	Q92fq1 pseudomonas	317	39	26.5	1528	4	Q95078	Q95078 homo sapien
245	39	26.5	125	12	Q72175	Q72175 hepatitis c	318	39	26.5	2018	5	Q20487	Q20487 caenorhabdi
246	39	26.5	131	13	Q91752	Q91752 xenopus lae	319	39	26.5	2160	5	Q17709	Q17709 caenorhabdi
247	39	26.5	140	11	Q9WU27	Q9wu27 cavia porce	320	39	26.5	5198	5	Q76518	Q76518 caenorhabdi
248	39	26.5	191	12	Q83020	Q83020 lactate deh	321	38.5	26.2	110	10	Q80873	Q80873 arabisopsis
249	39	26.5	191	12	Q06499	Q06499 lactate deh	322	38.5	26.2	118	5	Q44076	Q44076 caenorhabdi
250	39	26.5	191	12	Q9YS40	Q9ys40 lactate deh	323	38.5	26.2	153	2	Q84537	Q84537 chlamydia t
251	39	26.5	191	12	Q9YS35	Q9ys35 lactate deh	324	38.5	26.2	172	4	Q9Y6N0	Q9y6n0 homo sapien
252	39	26.5	213	2	P94356	P94356 bacillus su	325	38.5	26.2	209	1	Q20216	Q20216 caenorhabdi
253	39	26.5	217	13	Q91344	Q91344 gallus gall	326	38.5	26.2	211	1	Q30300	Q30300 archaeoglob
254	39	26.5	241	13	Q91017	Q91017 gallus gall	327	38.5	26.2	211	5	Q96305	Q96305 monoceromo
255	39	26.5	244	10	Q9XI70	Q9xi70 arabisopsis	328	38.5	26.2	260	13	Q9YGS6	Q9ygs6 brachydanio
256	39	26.5	246	2	P72704	P72704 synchocyst	329	38.5	26.2	284	5	Q25581	Q25581 osteragia
257	39	26.5	250	5	Q21496	Q21496 caenorhabdi	330	38.5	26.2	284	5	Q25582	Q25582 osteragia
258	39	26.5	258	5	Q9Y0E9	Q9y0e9 drosophila	331	38.5	26.2	320	2	Q68306	Q68306 nostoc pec7
259	39	26.5	276	2	Q54185	Q54185 streptococc	332	38.5	26.2	320	2	Q44215	Q44215 anabaena sp
260	39	26.5	276	12	Q86800	Q86800 tomato aspe	333	38.5	26.2	320	2	Q92AK3	Q92ak3 xenopus lae
261	39	26.5	285	4	Q94760	Q94760 homo sapien	334	38.5	26.2	340	13	Q91808	Q91808 xenopus lae
262	39	26.5	285	11	Q08557	Q08557 rattus norv	335	38.5	26.2	342	2	P94154	P94154 alcaligenes
263	39	26.5	291	5	Q76491	Q76491 brachiosteo	336	38.5	26.2	347	2	P95603	P95603 alcaligenes
264	39	26.5	300	4	Q9X3M3	Q9x3m3 homo sapien	337	38.5	26.2	347	13	Q91807	Q91807 xenopus lae
265	39	26.5	303	5	Q22981	Q22981 caenorhabdi	338	38.5	26.2	350	2	Q52644	Q52644 pseudomonas
266	39	26.5	316	5	Q926R2	Q926r2 chlamydia p	339	38.5	26.2	399	4	Q9Y6E8	Q9y6e8 homo sapien
267	39	26.5	319	2	Q28970	Q28970 archaeoglob	340	38.5	26.2	425	6	Q02661	Q02661 bos taurus
268	39	26.5	326	1	Q28970	Q28970 archaeoglob	341	38.5	26.2	504	2	Q92578	Q92578 streptomyce
269	39	26.5	349	13	Q9YHY7	Q9yhy7 gallus gall	342	38.5	26.2	504	4	Q75850	Q75850 homo sapien
270	39	26.5	365	2	P73132	P73132 synchocyst	343	38.5	26.2	508	5	Q16541	Q16541 caenorhabdi
271	39	26.5	367	2	Q67191	Q67191 aquifex aeo	344	38.5	26.2	577	5	Q18353	Q18353 drosophila
272	39	26.5	379	10	Q49316	Q49316 arabisopsis	345	38.5	26.2	586	5	Q18679	Q18679 caenorhabdi
273	39	26.5	380	8	Q48321	Q48321 hylobates h	346	38.5	26.2	647	3	Q93843	Q93843 emericeila
274	39	26.5	389	5	Q9XTD7	Q9xtd7 caenorhabdi	347	38.5	26.2	705	4	Q43377	Q43377 homo sapien
275	39	26.5	403	10	Q81844	Q81844 arabisopsis	348	38.5	26.2	715	5	Q21706	Q21706 caenorhabdi
276	39	26.5	406	4	Q13477	Q13477 homo sapien	349	38.5	26.2	863	4	Q9Y5C4	Q9y5c4 homo sapien
277	39	26.5	437	10	Q23024	Q23024 arabisopsis	350	38.5	26.2	904	4	Q08192	Q08192 homo sapien
278	39	26.5	439	3	Q92401	Q92401 agaricus bi	351	38.5	26.2	934	4	Q06022	Q06022 homo sapien
279	39	26.5	466	12	Q81016	Q81016 human papil	352	38.5	26.2	991	2	Q83528	Q83528 treponema p
280	39	26.5	475	2	P74540	P74540 synchocyst	353	38.5	26.2	1064	13	Q90601	Q90601 gallus gall
281	39	26.5	534	10	Q81790	Q81790 arabisopsis	354	38.5	26.2	1200	4	Q60247	Q60247 homo sapien
282	39	26.5	544	11	Q54802	Q54802 mus musculu	355	38.5	26.2	1840	13	Q90831	Q90831 gallus gall
283	39	26.5	553	2	Q50258	Q50258 micromonosp	356	38.5	26.2	4590	4	Q14517	Q14517 homo sapien
284	39	26.5	568	13	Q91572	Q91572 xenopus lae	357	38.5	26.2	4957	4	Q14687	Q14687 homo sapien
285	39	26.5	600	3	Q74977	Q74977 schizosacch	358	38.5	26.2	5262	4	Q14686	Q14686 homo sapien
286	39	26.5	643	2	Q9WZY8	Q9wzy8 thermotoga	359	38	25.9	39	4	Q14328	Q14328 homo sapien
287	39	26.5	655	1	Q27929	Q27929 methanobact	360	38	25.9	60	2	Q05893	Q05893 mycobacteri
288	39	26.5	658	1	Q27041	Q27041 methanobact	361	38	25.9	77	2	Q33036	Q33036 mycobacteri
289	39	26.5	678	4	Q94850	Q94850 homo sapien	362	38	25.9	87	4	Q16345	Q16345 homo sapien
290	39	26.5	691	2	Q55726	Q55726 synchocyst	363	38	25.9	106	5	Q96159	Q96159 plasmodium
291	39	26.5	697	4	Q43167	Q43167 homo sapien	364	38	25.9	121	13	Q91003	Q91003 gallus gall
292	39	26.5	729	1	Q26857	Q26857 methanobact	365	38	25.9	128	11	Q55178	Q55178 mus musculu
293	39	26.5	735	5	Q76912	Q76912 drosophila	366	38	25.9	137	5	Q18032	Q18032 caenorhabdi
294	39	26.5	754	4	Q9Y460	Q9y460 homo sapien	367	38	25.9	149	2	Q69501	Q69501 mycobacteri
295	39	26.5	770	4	Q75074	Q75074 homo sapien	368	38	25.9	166	3	Q9Y7N3	Q9y7n3 schizosacch
296	39	26.5	770	11	Q88204	Q88204 rattus norv	369	38	25.9	172	2	Q54789	Q54789 streptococc
297	39	26.5	780	10	Q22017	Q22017 cylindrothe	370	38	25.9	187	2	P71516	P71516 methylobact
298	39	26.5	794	10	Q92R11	Q92r11 arabisopsis	371	38	25.9	203	2	Q55581	Q55581 synchocyst
299	39	26.5	822	13	Q91288	Q91288 pleurodeles	372	38	25.9	209	12	Q81587	Q81587 hepatitis c
300	39	26.5	857	11	Q03717	Q03717 mus musculu	373	38	25.9	210	12	Q64848	Q64848 porcine ade
301	39	26.5	895	5	Q97228	Q97228 plasmodium	374	38	25.9	222	12	Q64901	Q64901 arctic grou
302	39	26.5	895	9	Q9ZXI8	Q9zx18 pseudomonas	375	38	25.9	228	11	Q08686	Q08686 mus musculu
303	39	26.5	986	10	P93416	P93416 oryza sativ	376	38	25.9	254	4	Q43518	Q43518 homo sapien
304	39	26.5	988	10	Q64454	Q64454 oryza sativ	377	38	25.9	256	12	Q81586	Q81586 hepatitis c
305	39	26.5	1012	3	Q94609	Q94609 schizosacch	378	38	25.9	263	4	Q00318	Q00318 homo sapien
306	39	26.5	1021	10	Q49548	Q49548 arabisopsis	379	38	25.9	267	10	Q82052	Q82052 sorghum bic
307	39	26.5	1131	10	Q93526	P93526 sorghum bic	380	38	25.9	267	10	Q82053	Q82053 sorghum bic
308	39	26.5	1139	2	Q50466	Q50466 mycobacteri	381	38	25.9	282	12	Q64900	Q64900 arctic grou
309	39	26.5	1156	5	Q21214	Q21214 caenorhabdi	382	38	25.9	286	2	Q30544	Q30544 agrobacteri
310	39	26.5	1238	4	Q95289	Q95289 homo sapien	383	38	25.9	287	13	Q9W7B8	Q9w7b8 brachydanio
311	39	26.5	1298	4	Q99881	Q99881 homo sapien	384	38	25.9	297	11	Q9Z1S9	Q9z1s9 cavia porce

385	38	25.9	298	10	P93790	P93790 triticum ae	458	38	25.9	1021	5	Q20217	Q20217 caenorhabdi
386	38	25.9	299	2	O06040	O06040 lactococcus	459	38	25.9	1072	5	Q26157	Q26157 plasmodium
387	38	25.9	300	5	O76409	O76409 caenorhabdi	460	38	25.9	1087	5	Q26156	Q26156 plasmodium
388	38	25.9	314	10	O23763	O23763 cucurbita m	461	38	25.9	1089	5	Q26155	Q26155 plasmodium
389	38	25.9	319	5	O22676	O22676 caenorhabdi	462	38	25.9	1263	1	P81409	P81409 pyrococcus
390	38	25.9	325	5	O9XTR8	O9XTR8 caenorhabdi	463	38	25.9	1308	5	O96129	O96129 plasmodium
391	38	25.9	327	1	O9YA75	O9YA75 aeropyrum p	464	38	25.9	1323	3	O14228	O14228 schizosacch
392	38	25.9	330	5	P90535	P90535 dictyosteli	465	38	25.9	1346	2	O29612	O29612 streptomyc
393	38	25.9	341	3	O9W788	O9W788 schizosacch	466	38	25.9	1369	12	Q64907	Q64907 alcelaphine
394	38	25.9	343	2	Q53025	Q53025 nocardia co	467	38	25.9	1581	13	O73809	O73809 fugu rubrip
395	38	25.9	353	5	Q23166	Q23166 caenorhabdi	468	38	25.9	1616	4	O15054	O15054 homo sapien
396	38	25.9	367	5	O26723	O26723 trypanosoma	469	38	25.9	2168	5	Q19330	Q19330 caenorhabdi
397	38	25.9	369	4	Q06246	Q06246 homo sapien	470	38	25.9	2606	12	O36414	O36414 alcelaphine
398	38	25.9	375	11	O88701	O88701 mus musculu	471	38	25.9	2843	4	O9Y6R7	O9Y6R7 homo sapien
399	38	25.9	381	5	O01861	O01861 caenorhabdi	472	38	25.9	2864	12	Q69422	Q69422 hepatitis g
400	38	25.9	404	5	Q20044	Q20044 caenorhabdi	473	38	25.9	2948	5	Q93785	Q93785 caenorhabdi
401	38	25.9	409	5	Q23046	Q23046 caenorhabdi	474	38	25.9	3037	12	Q68749	Q68749 hepatitis c
402	38	25.9	413	2	O07310	O07310 rhodobacter	475	38	25.9	4123	4	O75851	O75851 homo sapien
403	38	25.9	413	13	O98860	O98860 cynops pyrr	476	37.5	25.5	162	12	O89900	O89900 human herpe
404	38	25.9	427	12	Q64899	Q64899 arctic grou	477	37.5	25.5	206	10	O9XIS8	O9XIS8 glycine max
405	38	25.9	436	2	Q45144	Q45144 corynebacte	478	37.5	25.5	290	5	O44173	O44173 caenorhabdi
406	38	25.9	438	5	O9XIS1	O9XIS1 ciona intes	479	37.5	25.5	291	4	Q92567	Q92567 homo sapien
407	38	25.9	444	11	Q62219	Q62219 mus musculu	480	37.5	25.5	299	3	Q99344	Q99344 saccharomyc
408	38	25.9	452	2	O86909	O86909 sphingomona	481	37.5	25.5	303	2	P73529	P73529 synechocyst
409	38	25.9	465	12	O92450	O92450 bombyx mori	482	37.5	25.5	321	12	Q66648	Q66648 equine herp
410	38	25.9	475	12	O81588	O81588 hepatitis c	483	37.5	25.5	357	10	O43639	O43639 secate cere
411	38	25.9	485	4	Q13434	Q13434 homo sapien	484	37.5	25.5	357	10	O04365	O04365 secate cere
412	38	25.9	489	2	Q9X7P0	Q9X7P0 streptomyc	485	37.5	25.5	367	3	O01471	O01471 paecilomyc
413	38	25.9	504	11	O35841	O35841 mus musculu	486	37.5	25.5	371	2	O06555	O06555 mycobacteri
414	38	25.9	507	5	Q94551	Q94551 dugesia tig	487	37.5	25.5	374	13	O73712	O73712 brachydanio
415	38	25.9	508	4	Q02930	Q02930 homo sapien	488	37.5	25.5	374	13	O73715	O73715 brachydanio
416	38	25.9	545	12	P88958	P88958 kaposi's sa	489	37.5	25.5	375	4	O43494	O43494 homo sapien
417	38	25.9	559	13	O9YXG5	O9YXG5 brachydanio	490	37.5	25.5	379	3	O9Y759	O9Y759 beauveria b
418	38	25.9	588	5	Q22747	Q22747 caenorhabdi	491	37.5	25.5	391	4	Q16226	Q16226 homo sapien
419	38	25.9	594	5	Q24970	Q24970 giardia lam	492	37.5	25.5	411	13	O42252	O42252 gallus gall
420	38	25.9	599	10	O41428	O41428 solanum tub	493	37.5	25.5	421	5	O9XW36	O9XW36 caenorhabdi
421	38	25.9	600	13	O93386	O93386 brachydanio	494	37.5	25.5	449	1	O27341	O27341 methanobact
422	38	25.9	606	5	O44940	O44940 drosophila	495	37.5	25.5	453	5	Q24771	Q24771 elmeria ace
423	38	25.9	606	5	O46104	O46104 drosophila	496	37.5	25.5	476	3	Q12186	Q12186 saccharomyc
424	38	25.9	622	12	Q9YNA0	Q9YNA0 sheep pulmo	497	37.5	25.5	477	12	Q96800	Q96800 friend mink
425	38	25.9	622	12	O9YJ35	O9YJ35 sheep pulmo	498	37.5	25.5	493	4	O15048	O15048 homo sapien
426	38	25.9	622	12	O9WJR2	O9WJR2 sheep pulmo	499	37.5	25.5	506	3	O00863	O00863 nectria bae
427	38	25.9	645	3	Q42946	Q42946 schizosacch	500	37.5	25.5	514	5	P90547	P90547 entamoeba i
428	38	25.9	650	1	O9YB17	O9YB17 aeropyrum p	501	37.5	25.5	522	4	O9Y4K3	O9Y4K3 homo sapien
429	38	25.9	650	5	O96380	O96380 spizula sol	502	37.5	25.5	541	10	O65305	O65305 hordeum vul
430	38	25.9	675	2	Q9ZCF6	Q9ZCF6 rickettsia	503	37.5	25.5	544	2	Q60027	Q60027 thymomonos
431	38	25.9	678	5	O18106	O18106 caenorhabdi	504	37.5	25.5	576	2	O50865	O50865 myxococcus
432	38	25.9	685	5	O44249	O44249 manduca sex	505	37.5	25.5	669	12	O9YYS3	O9YYS3 murine leuk
433	38	25.9	685	5	Q27451	Q27451 bombyx mori	506	37.5	25.5	730	2	Q9X6U0	Q9X6U0 streptococc
434	38	25.9	692	2	O53767	O53767 mycobacteri	507	37.5	25.5	737	5	Q23985	Q23985 drosophila
435	38	25.9	747	13	O91900	O91900 xenopus lae	508	37.5	25.5	772	4	O75676	O75676 homo sapien
436	38	25.9	748	11	O62177	O62177 mus musculu	509	37.5	25.5	789	11	Q63180	Q63180 rattus norv
437	38	25.9	757	12	O37361	O37361 carrot red	510	37.5	25.5	830	10	O9ZT06	O9ZT06 arabidopsis
438	38	25.9	766	1	O58222	O58222 pyrococcus	511	37.5	25.5	906	10	O9XIF0	O9XIF0 arabidopsis
439	38	25.9	769	5	Q24571	Q24571 giardia lam	512	37.5	25.5	1035	13	O57537	O57537 xenopus lae
440	38	25.9	771	4	Q14563	Q14563 homo sapien	513	37.5	25.5	1069	4	O60245	O60245 mus sapien
441	38	25.9	772	11	O08665	O08665 mus musculu	514	37.5	25.5	1069	11	O88185	O88185 mus musculu
442	38	25.9	772	11	Q63548	Q63548 rattus norv	515	37.5	25.5	1072	4	O60246	O60246 homo sapien
443	38	25.9	772	13	O90607	O90607 gallus gall	516	37.5	25.5	1127	12	O70652	O70652 gibbon ape
444	38	25.9	778	13	Q9W686	Q9W686 brachydanio	517	37.5	25.5	1171	13	O42094	O42094 gallus gall
445	38	25.9	812	5	O21435	O21435 caenorhabdi	518	37.5	25.5	1197	12	O69101	O69101 herpes simp
446	38	25.9	813	10	Q9ZSN3	Q9ZSN3 phaseolus v	519	37.5	25.5	1218	11	O70365	O70365 mus musculu
447	38	25.9	860	10	O9X113	O9X113 arabidopsis	520	37.5	25.5	1231	5	Q26153	Q26153 plasmodium
448	38	25.9	876	13	Q91593	Q91593 xenopus lae	521	37.5	25.5	1581	4	O15065	O15065 homo sapien
449	38	25.9	884	4	O9Y6R5	O9Y6R5 homo sapien	522	37.5	25.5	1587	4	O15741	O15741 homo sapien
450	38	25.9	917	2	O9ZMQ1	O9ZMQ1 pseudomonas	523	37.5	25.5	1588	4	O15293	O15293 homo sapien
451	38	25.9	925	10	Q9ZTA5	Q9ZTA5 arabidopsis	524	37.5	25.5	1602	11	O08873	O08873 rattus norv
452	38	25.9	939	2	O32491	O32491 bacteroides	525	37.5	25.5	2590	13	Q9W7R4	Q9W7R4 brachydanio
453	38	25.9	963	5	O44393	O44393 pisaster oc	526	37.5	25.5	2590	13	Q63200	Q63200 rattus norv
454	38	25.9	964	10	O9ZM02	O9ZM02 arabidopsis	527	37.5	25.2	77	5	O77416	O77416 anisakis si
455	38	25.9	968	10	O9ZM07	O9ZM07 arabidopsis	528	37.5	25.2	84	2	P73792	P73792 synechocyst
456	38	25.9	972	10	O9ZW06	O9ZW06 arabidopsis	529	37.5	25.2	86	2	O54474	O54474 staphylococ
457	38	25.9	1017	5	O44728	O44728 caenorhabdi	530	37.5	25.2	94	4	Q14327	Q14327 homo sapien

531	37	25.2	94	12	Q68354	Q68354 hepatitis c	604	37	25.2	706	13	O42593	O42593 xenopus lae
532	37	25.2	160	2	Q83022	O83022 rhodobacter	605	37	25.2	708	13	P87363	P87363 gallus gall
533	37	25.2	161	2	O46226	O46226 chlamydia p	606	37	25.2	715	10	Q92VE2	Q92VE2 arabidopsis
534	37	25.2	169	5	Q9XW8	Q9XW8 caenorhabdi	607	37	25.2	728	11	P97350	P97350 mus musculu
535	37	25.2	179	13	Q9W7M2	Q9W7M2 cryzias lat	608	37	25.2	749	4	Q13214	Q13214 homo sapien
536	37	25.2	188	12	Q90386	Q90386 rhesus cyto	609	37	25.2	783	5	P92163	P92163 strongyloce
537	37	25.2	208	5	Q94164	Q94164 caenorhabdi	610	37	25.2	783	13	Q90275	Q90275 brachydanio
538	37	25.2	221	12	Q64824	Q64824 human adeno	611	37	25.2	790	13	Q91838	Q91838 xenopus lae
539	37	25.2	228	10	Q40274	Q40274 malus domes	612	37	25.2	801	5	P91774	P91774 pacifastacu
540	37	25.2	231	5	Q15714	Q15714 dictyosteli	613	37	25.2	803	3	Q13657	Q13657 schizosacch
541	37	25.2	267	13	Q9W694	Q9W694 gallus gall	614	37	25.2	822	2	Q07806	Q07806 pseudomonas
542	37	25.2	281	2	O52811	O52811 amycolatops	615	37	25.2	860	13	Q9W7J1	Q9W7J1 brachydanio
543	37	25.2	305	4	Q15272	Q15272 homo sapien	616	37	25.2	871	2	Q9XDM3	Q9XDM3 aquifex pyr
544	37	25.2	325	2	Q92889	Q92889 chlamydia p	617	37	25.2	917	13	Q98931	Q98931 gallus gall
545	37	25.2	330	11	P70271	P70271 mus musculu	618	37	25.2	927	11	Q62402	Q62402 mus musculu
546	37	25.2	334	4	Q75829	Q75829 homo sapien	619	37	25.2	935	4	Q13078	Q13078 homo sapien
547	37	25.2	339	12	Q69306	Q69306 gallid herp	620	37	25.2	946	5	O96163	O96163 plasmodium
548	37	25.2	346	5	O02240	O02240 caenorhabdi	621	37	25.2	950	5	O18367	O18367 drosophila
549	37	25.2	351	2	Q4024	Q4024 chlamydia p	622	37	25.2	950	5	O18367	O18367 pneumocysti
550	37	25.2	355	2	Q67508	Q67508 aquifex aeo	623	37	25.2	1026	3	Q74669	Q74669 pneumocysti
551	37	25.2	365	11	Q62150	Q62150 mus musculu	624	37	25.2	1031	12	Q03261	Q03261 chimpanzee
552	37	25.2	369	12	Q68869	Q68869 hepatitis c	625	37	25.2	1034	11	O35888	O35888 rattus norv
553	37	25.2	373	2	Q9X800	Q9X800 streptomyc	626	37	25.2	1065	10	O48948	O48948 arabidopsis
554	37	25.2	379	10	O23819	O23819 arabis gemm	627	37	25.2	1072	3	Q94537	Q94537 schizosacch
555	37	25.2	380	8	Q47890	Q47890 hylobates g	628	37	25.2	1077	11	Q62217	Q62217 mus musculu
556	37	25.2	380	8	Q47892	Q47892 hylobates l	629	37	25.2	1148	4	Q99590	Q99590 homo sapien
557	37	25.2	380	8	Q47894	Q47894 hylobates s	630	37	25.2	1166	4	Q9Y213	Q9Y213 homo sapien
558	37	25.2	380	8	Q47895	Q47895 hylobates s	631	37	25.2	1182	4	Q99495	Q99495 homo sapien
559	37	25.2	384	12	O82000	O82000 human papil	632	37	25.2	1186	5	Q17786	Q17786 caenorhabdi
560	37	25.2	385	5	Q93787	Q93787 caenorhabdi	633	37	25.2	1217	2	O45440	Q45440 bacillus sp
561	37	25.2	386	10	Q49617	Q49617 arabidopsis	634	37	25.2	1220	2	Q9XDH5	Q9XDH5 thermus aqu
562	37	25.2	389	1	O34135	O34135 halobacteri	635	37	25.2	1238	12	Q72768	Q72768 eastern equ
563	37	25.2	399	10	Q41241	Q41241 lycopersico	636	37	25.2	1241	12	O66579	O66579 eastern equ
564	37	25.2	390	10	Q41242	Q41242 lycopersico	637	37	25.2	1242	12	O88678	O88678 eastern equ
565	37	25.2	395	2	O67558	O67558 aquifex aeo	638	37	25.2	1242	12	O88790	O88790 eastern equ
566	37	25.2	396	5	Q04151	Q04151 toxoplasma	639	37	25.2	1242	12	O88792	O88792 eastern equ
567	37	25.2	398	12	Q67633	Q67633 gallid herp	640	37	25.2	1242	12	O88793	O88793 eastern equ
568	37	25.2	403	6	Q28410	Q28410 felis silve	641	37	25.2	1242	12	O88794	O88794 eastern equ
569	37	25.2	405	1	Q9YE99	Q9YE99 aeropyrum p	642	37	25.2	1242	12	O88795	O88795 eastern equ
570	37	25.2	406	2	P95619	P95619 rhodocyclu	643	37	25.2	1242	12	O88796	O88796 eastern equ
571	37	25.2	425	6	O28759	O28759 procavia ca	644	37	25.2	1242	12	O88797	O88797 eastern equ
572	37	25.2	432	11	Q64385	Q64385 mus musculu	645	37	25.2	1242	12	O88798	O88798 eastern equ
573	37	25.2	432	11	P70225	P70225 mus musculu	646	37	25.2	1242	12	O88799	O88799 eastern equ
574	37	25.2	438	2	O66179	O66179 agrobacteri	647	37	25.2	1242	12	O08359	O08359 eastern equ
575	37	25.2	441	11	O88598	O88598 mus musculu	648	37	25.2	1243	4	O43314	O43314 homo sapien
576	37	25.2	442	4	Q76088	Q76088 homo sapien	649	37	25.2	1262	11	Q60988	Q60988 mus musculu
577	37	25.2	444	4	Q9Y2V5	Q9Y2V5 homo sapien	650	37	25.2	1265	4	O94899	O94899 homo sapien
578	37	25.2	446	6	O19071	O19071 s alpha-1,6	651	37	25.2	1287	4	O15468	O15468 homo sapien
579	37	25.2	460	4	O43294	O43294 homo sapien	652	37	25.2	1313	5	O01795	O01795 caenorhabdi
580	37	25.2	472	10	Q92RQ0	Q92RQ0 arabidopsis	653	37	25.2	1332	5	O45599	O45599 caenorhabdi
581	37	25.2	485	2	Q9X623	Q9X623 bacillus sp	654	37	25.2	1344	11	O35851	O35851 mus musculu
582	37	25.2	489	2	P72455	P72455 streptomyc	655	37	25.2	1375	13	Q9YGM5	Q9YGM5 fugu rubrip
583	37	25.2	492	10	O24511	O24511 nicotiana t	656	37	25.2	1437	12	Q9WF77	Q9WF77 human immun
584	37	25.2	504	4	Q9Y4J7	Q9Y4J7 homo sapien	657	37	25.2	1462	11	O61026	O61026 mus musculu
585	37	25.2	505	2	O54393	O54393 myxococcus	658	37	25.2	1506	5	O61213	O61213 caenorhabdi
586	37	25.2	515	2	P96182	P96182 wolinnella s	659	37	25.2	1612	11	O89026	O89026 mus musculu
587	37	25.2	517	12	Q9YU51	Q9YU51 hemorrhagic	660	37	25.2	1651	11	O55005	O55005 rattus norv
588	37	25.2	522	13	Q9YGM4	Q9YGM4 brachydanio	661	37	25.2	1716	4	O14528	O14528 homo sapien
589	37	25.2	523	10	Q41219	Q41219 glycine max	662	37	25.2	1734	11	O60592	O60592 mus musculu
590	37	25.2	528	4	O15441	O15441 homo sapien	663	37	25.2	1921	5	O01349	O01349 drosophila
591	37	25.2	528	5	Q9XU53	Q9XU53 caenorhabdi	664	37	25.2	2126	2	O07798	O07798 mycobacteri
592	37	25.2	541	5	Q17941	Q17941 caenorhabdi	665	37	25.2	2137	4	O15021	O15021 homo sapien
593	37	25.2	546	5	O17942	O17942 caenorhabdi	666	37	25.2	2334	12	Q96725	Q96725 european br
594	37	25.2	571	3	P78802	P78802 schizosacch	667	37	25.2	2358	3	O94638	O94638 schizosacch
595	37	25.2	580	4	O00456	O00456 homo sapien	668	37	25.2	2605	2	Q50858	Q50858 myxococcus
596	37	25.2	593	2	O83179	O83179 treponema p	669	37	25.2	3006	5	O26032	O26032 plasmodium
597	37	25.2	597	5	Q07317	Q07317 giardia int	670	37	25.2	3010	12	P88803	P88803 hepatitis c
598	37	25.2	605	5	O45302	O45302 caenorhabdi	671	37	25.2	3110	4	Q93022	Q93022 homo sapien
599	37	25.2	609	5	O17122	O17122 caenorhabdi	672	37	25.2	3164	12	O69088	O69088 human herpe
600	37	25.2	614	11	O88561	O88561 mus musculu	673	37	25.2	3675	6	Q28733	Q28733 oryctolagus
601	37	25.2	635	3	P87234	P87234 schizosacch	674	37	25.2	10797	3	O94116	O94116 aureobasidi
602	37	25.2	648	4	P78425	P78425 homo sapien	675	37	25.2	36.5	2	Q92ND7	Q92nd7 thermus aqu
603	37	25.2	684	4	P78424	P78424 homo sapien	676	37	25.2	24.8	54		

677	36.5	24.8	131	8	O21239	O21239 reclinomona	750	36	24.5	180	4	Q99942	Q99942 homo sapien
678	36.5	24.8	165	10	Q43136	Q43136 sorghum bic	751	36	24.5	182	2	Q9x620	Q9x620 bordetella
679	36.5	24.8	209	2	P71652	P71652 mycobacteri	752	36	24.5	184	5	Q23830	Q23830 cryptospori
680	36.5	24.8	214	5	Q21830	Q21830 caenorhabdi	753	36	24.5	185	4	Q95300	Q95300 homo sapien
681	36.5	24.8	222	2	P73373	P73373 synecocyst	754	36	24.5	193	5	Q18282	Q18282 caenorhabdi
682	36.5	24.8	249	5	O45764	O45764 caenorhabdi	755	36	24.5	199	5	P91464	P91464 caenorhabdi
683	36.5	24.8	261	3	O74801	O74801 schizosacch	756	36	24.5	201	12	Q82955	Q82955 kaposi's sa
684	36.5	24.8	269	11	Q92205	Q92205 mus musculu	757	36	24.5	215	2	Q55934	Q55934 synecocyst
685	36.5	24.8	273	11	Q60742	Q60742 mus musculu	758	36	24.5	217	11	P97833	P97833 rattus norv
686	36.5	24.8	279	11	Q60740	Q60740 mus musculu	759	36	24.5	219	3	O60186	O60186 schizosacch
687	36.5	24.8	299	10	O49201	O49201 gossypium h	760	36	24.5	221	6	Q28090	Q28090 bos taurus
688	36.5	24.8	308	11	Q60743	Q60743 mus musculu	761	36	24.5	221	6	O97631	O97631 oviss aries
689	36.5	24.8	316	6	Q9XSC7	Q9XSC7 sus scrofa	762	36	24.5	223	6	O9XTAL	O9XTAL felis silve
690	36.5	24.8	317	6	O97599	O97599 bos taurus	763	36	24.5	223	6	O9XSV7	O9XSV7 felis silve
691	36.5	24.8	321	5	Q18199	Q18199 caenorhabdi	764	36	24.5	223	6	Q9XS11	Q9XS11 canis famill
692	36.5	24.8	328	2	Q92JW2	Q92JW2 helicobacte	765	36	24.5	223	11	Q62859	Q62859 rattus norv
693	36.5	24.8	360	2	O53282	O53282 mycobacteri	766	36	24.5	224	11	Q920L2	Q920L2 mus musculu
694	36.5	24.8	365	4	O13777	O13777 homo sapien	767	36	24.5	224	12	Q69112	Q69112 herpes simp
695	36.5	24.8	378	5	Q27383	Q27383 caenorhabdi	768	36	24.5	226	12	O86833	O86833 moloney mur
696	36.5	24.8	388	5	O44606	O44606 caenorhabdi	769	36	24.5	236	4	Q9V4X9	Q9V4X9 homo sapien
697	36.5	24.8	425	13	Q91867	Q91867 xenopus lae	770	36	24.5	243	1	O29719	O29719 archaeoglob
698	36.5	24.8	437	13	O13111	O13111 gallus gall	771	36	24.5	244	2	P78271	P78271 escherichia
699	36.5	24.8	459	10	O82718	O82718 lupinus lut	772	36	24.5	246	2	P94338	P94338 corynebacte
700	36.5	24.8	463	2	O50190	Q50190 mycobacteri	773	36	24.5	246	2	O24747	O24747 corynebacte
701	36.5	24.8	465	10	Q9ZU93	Q9ZU93 arabidopsis	774	36	24.5	247	12	O86834	O86834 moloney mur
702	36.5	24.8	526	11	O08783	O08783 mus musculu	775	36	24.5	253	1	O27844	O27844 methanobact
703	36.5	24.8	534	2	Q92353	Q92353 synecococc	776	36	24.5	253	11	O08602	O08602 mus musculu
704	36.5	24.8	548	2	Q59317	Q59317 synecocyst	777	36	24.5	257	12	P88391	P88391 human immun
705	36.5	24.8	550	2	P73918	P73918 synecocyst	778	36	24.5	257	12	P88397	P88397 human immun
706	36.5	24.8	551	3	O08109	O08109 saccharomyc	779	36	24.5	260	10	O81279	O81279 glycine max
707	36.5	24.8	553	4	O14806	O14806 homo sapien	780	36	24.5	260	11	Q92334	Q92334 rattus norv
708	36.5	24.8	561	5	P90667	P90667 aplysia cal	781	36	24.5	261	10	O41551	O41551 triticum ae
709	36.5	24.8	561	5	P90668	P90668 aplysia cal	782	36	24.5	264	2	O85476	O85476 escherichia
710	36.5	24.8	561	5	P92190	P92190 aplysia cal	783	36	24.5	267	4	Q15817	Q15817 homo sapien
711	36.5	24.8	570	10	O81140	O81140 populus tre	784	36	24.5	276	10	O41552	O41552 triticum ae
712	36.5	24.8	580	3	Q94547	Q94547 schizosacch	785	36	24.5	277	10	Q39602	Q39602 chlamydomon
713	36.5	24.8	637	4	O94878	O94878 homo sapien	786	36	24.5	277	11	O06477	O06477 mus musculu
714	36.5	24.8	639	2	P77931	P77931 pseudomonas	787	36	24.5	281	12	Q9WD92	Q9WD92 influenza b
715	36.5	24.8	656	4	O76103	O76103 homo sapien	788	36	24.5	281	12	Q9WD87	Q9WD87 influenza b
716	36.5	24.8	736	5	O01441	O01441 caenorhabdi	789	36	24.5	281	12	Q9WD85	Q9WD85 influenza b
717	36.5	24.8	754	5	P91063	P91063 caenorhabdi	790	36	24.5	286	2	O31581	O31581 bacillus su
718	36.5	24.8	821	6	O19060	O19060 saguinus oe	791	36	24.5	286	10	O82358	O82358 arabidopsis
719	36.5	24.8	832	3	O74203	O74203 phanerochae	792	36	24.5	287	4	Q03969	Q03969 homo sapien
720	36.5	24.8	833	5	O16461	O16461 caenorhabdi	793	36	24.5	289	1	Q9VBP2	Q9VBP2 aeropyrum p
721	36.5	24.8	865	1	O27595	O27595 methanobact	794	36	24.5	292	4	O00448	O00448 homo sapien
722	36.5	24.8	965	5	O44405	O44405 caenorhabdi	795	36	24.5	300	12	Q9WR44	Q9WR44 cercopithe
723	36.5	24.8	1019	5	O96435	O96435 drosophila	796	36	24.5	301	6	O19055	O19055 papio hamad
724	36.5	24.8	1035	5	O9XZ16	Q9XZ16 drosophila	797	36	24.5	305	11	O70184	O70184 cavia porce
725	36.5	24.8	1059	5	P90884	P90884 caenorhabdi	798	36	24.5	306	3	P78583	P78583 aspergillus
726	36.5	24.8	1075	2	O06342	O06342 mycobacteri	799	36	24.5	313	10	O41546	O41546 triticum ae
727	36.5	24.8	1080	5	Q9Y1J0	Q9Y1J0 drosophila	800	36	24.5	317	1	P77944	P77944 sulfolobus
728	36.5	24.8	1405	11	O61202	O61202 mus musculu	801	36	24.5	322	5	O20218	O20218 caenorhabdi
729	36.5	24.8	1405	11	P70366	P70366 mus musculu	802	36	24.5	329	5	O44759	O44759 caenorhabdi
730	36.5	24.8	1410	2	O52673	O52673 escherichia	803	36	24.5	331	10	O9XGC6	O9XGC6 zea mays (m
731	36.5	24.8	1447	11	P70365	P70365 mus musculu	804	36	24.5	333	5	Q25689	Q25689 pseudoterra
732	36.5	24.8	1547	5	Q26471	Q26471 schistocerc	805	36	24.5	335	2	O30387	O30387 myxococcus
733	36.5	24.8	1610	3	O74349	O74349 schizosacch	806	36	24.5	341	10	Q9ZT38	Q9ZT38 glycine max
734	36.5	24.8	2054	11	O55164	O55164 rattus norv	807	36	24.5	343	4	O14685	O14685 homo sapien
735	36.5	24.8	2523	2	O53393	O53393 mycobacteri	808	36	24.5	343	4	O95161	O95161 homo sapien
736	36.5	24.8	9376	2	O85168	O85168 pseudomonas	809	36	24.5	345	8	O48029	O48029 charina bot
737	36	24.5	25	8	Q33292	Q33292 zea mays (m	810	36	24.5	350	10	O49958	O49958 triticum du
738	36	24.5	43	10	Q09097	Q09097 avena sativ	811	36	24.5	351	2	Q55017	Q55017 synecococc
739	36	24.5	75	12	O85323	O85323 vaccinia vi	812	36	24.5	352	4	O60426	O60426 homo sapien
740	36	24.5	83	2	P73480	P73480 synecocyst	813	36	24.5	354	5	O17034	O17034 caenorhabdi
741	36	24.5	103	12	O68355	O68355 hepatitis c	814	36	24.5	359	12	O85453	O85453 murine sarc
742	36	24.5	131	3	O08045	O08045 saccharomyc	815	36	24.5	361	5	O24397	O24397 drosophila
743	36	24.5	134	2	Q9WWN5	Q9WWN5 pseudomonas	816	36	24.5	362	4	O13236	O13236 homo sapien
744	36	24.5	136	6	Q95295	Q95295 sus scrofa	817	36	24.5	367	5	Q22069	Q22069 caenorhabdi
745	36	24.5	138	2	P73084	P73084 synecocyst	818	36	24.5	374	5	O01584	O01584 caenorhabdi
746	36	24.5	146	2	P96183	P96183 wolinnella s	819	36	24.5	381	2	O48919	O48919 mycobacteri
747	36	24.5	165	1	O29957	O29957 archaeoglob	820	36	24.5	384	3	O42707	O42707 schizosacch
748	36	24.5	165	1	O29066	O29066 archaeoglob	821	36	24.5	385	2	O84765	O84765 chlamydia t
749	36	24.5	179	1	O30033	O30033 archaeoglob	822	36	24.5	385	2	Q9Z703	Q9Z703 chlamydia p

823	36	24.5	392	2	P96214	P96214 mycobacteri	896	36	24.5	782	11	Q62445	Q62445 mus musculus
824	36	24.5	393	2	O82996	O82996 aeromonas c	897	36	24.5	783	11	Q63158	Q63158 rattus norv
825	36	24.5	400	1	O30076	O30076 archaeoglob	898	36	24.5	784	4	O75766	O75766 homo sapien
826	36	24.5	403	4	O75998	O75998 homo sapien	899	36	24.5	785	12	O93119	O93119 vaccinia vi
827	36	24.5	405	3	Q99079	Q99079 ustilago ma	900	36	24.5	793	4	O14968	O14968 homo sapien
828	36	24.5	411	5	P91419	P91419 caenorhabdi	901	36	24.5	796	2	Q924W5	Q924W5 streptomyce
829	36	24.5	413	11	P97343	P97343 mus musculus	902	36	24.5	840	2	Q9X8B0	Q9X8B0 streptomyce
830	36	24.5	414	6	Q28904	Q28904 sus scrofa	903	36	24.5	841	2	Q9X5V3	Q9X5V3 rhizobium l
831	36	24.5	416	6	O18936	O18936 amblysomus	904	36	24.5	842	11	Q61137	Q61137 mus musculus
832	36	24.5	419	11	Q63285	Q63285 rattus norv	905	36	24.5	888	4	O94880	O94880 homo sapien
833	36	24.5	421	4	Q15278	Q15278 homo sapien	906	36	24.5	890	11	Q921A0	Q921A0 cavia porce
834	36	24.5	425	5	Q15755	Q15755 dictyosteli	907	36	24.5	898	13	Q91592	Q91592 xenopus lae
835	36	24.5	427	6	Q28352	Q28352 dugong dugo	908	36	24.5	900	10	Q92SN2	Q92SN2 phaseolus v
836	36	24.5	439	13	Q42464	Q42464 salmo salar	909	36	24.5	919	10	O24375	O24375 solanum tub
837	36	24.5	461	3	Q92226	Q92226 emericella	910	36	24.5	924	5	Q20189	Q20189 caenorhabdi
838	36	24.5	462	1	O29558	O29558 archaeoglob	911	36	24.5	957	4	O14651	O14651 homo sapien
839	36	24.5	463	11	Q9WUP2	Q9WUP2 mus musculus	912	36	24.5	963	4	Q14114	Q14114 homo sapien
840	36	24.5	467	13	Q42465	Q42465 salmo salar	913	36	24.5	988	11	P97432	P97432 mus musculus
841	36	24.5	468	4	O95623	O95623 homo sapien	914	36	24.5	992	11	Q9WU50	Q9WU50 mus musculus
842	36	24.5	470	5	Q17317	Q17317 ceratitidis c	915	36	24.5	1014	5	Q26152	Q26152 plasmodium
843	36	24.5	490	11	P97443	P97443 mus musculus	916	36	24.5	1015	5	Q09938	Q09938 caenorhabdi
844	36	24.5	492	10	Q9XHH3	Q9XHH3 lycopersico	917	36	24.5	1017	2	O06586	O06586 mycobacteri
845	36	24.5	495	5	O45434	O45434 caenorhabdi	918	36	24.5	1063	10	Q9XEJ4	Q9XEJ4 zea mays (m
846	36	24.5	496	2	Q9XON9	Q9XON9 thermotoga	919	36	24.5	1083	2	O86637	O86637 streptomyce
847	36	24.5	497	13	Q73804	Q73804 fugu rubrip	920	36	24.5	1084	12	O65029	O65029 arabis mosa
848	36	24.5	498	11	Q62921	Q62921 rattus norv	921	36	24.5	1100	5	O96166	O96166 plasmodium
849	36	24.5	498	11	Q9WUB0	Q9WUB0 mus musculus	922	36	24.5	1108	10	Q9XIE1	Q9XIE1 arabidopsis
850	36	24.5	500	10	O81413	O81413 glycine max	923	36	24.5	1142	4	O60732	O60732 homo sapien
851	36	24.5	505	1	Q59015	Q59015 methanococc	924	36	24.5	1142	4	O75451	O75451 homo sapien
852	36	24.5	506	2	Q92513	Q92513 mycobacteri	925	36	24.5	1172	4	O9Y4F2	O9Y4F2 homo sapien
853	36	24.5	509	12	Q9YQ03	Q9YQ03 myxoma viru	926	36	24.5	1186	11	O08961	O08961 rattus norv
854	36	24.5	512	5	O76725	O76725 caenorhabdi	927	36	24.5	1206	4	O94860	O94860 homo sapien
855	36	24.5	513	11	Q9Z2W7	Q9Z2W7 rattus norv	928	36	24.5	1215	10	Q9ZVE1	Q9ZVE1 arabidopsis
856	36	24.5	514	5	Q93558	Q93558 caenorhabdi	929	36	24.5	1221	11	Q9WUI2	Q9WUI2 mus musculus
857	36	24.5	518	5	O01465	O01465 caenorhabdi	930	36	24.5	1231	11	O61484	O61484 mus musculus
858	36	24.5	525	3	O13389	O13389 emericella	931	36	24.5	1302	4	O14525	O14525 homo sapien
859	36	24.5	530	5	O16485	O16485 caenorhabdi	932	36	24.5	1376	12	P88917	P88917 kaposi's sa
860	36	24.5	535	13	O57405	O57405 gallus gall	933	36	24.5	1376	12	O40925	O40925 kaposi's sa
861	36	24.5	536	3	O13786	O13786 schizosacch	934	36	24.5	1376	12	Q98139	Q98139 kaposi's sa
862	36	24.5	537	2	O46977	O46977 escherichia	935	36	24.5	1464	4	O15596	O15596 homo sapien
863	36	24.5	542	5	Q23446	Q23446 caenorhabdi	936	36	24.5	1465	11	Q9WUI9	Q9WUI9 rattus norv
864	36	24.5	546	10	O04034	O04034 arabidopsis	937	36	24.5	1469	2	Q9X6V1	Q9X6V1 aquifex pyr
865	36	24.5	555	4	Q12912	Q12912 homo sapien	938	36	24.5	1475	5	O18647	O18647 caenorhabdi
866	36	24.5	561	11	P70166	P70166 mus musculus	939	36	24.5	1545	11	O08649	O08649 mus musculus
867	36	24.5	566	3	Q06696	Q06696 saccharomyc	940	36	24.5	1597	11	O08648	O08648 mus musculus
868	36	24.5	574	4	Q9Y5V3	Q9Y5V3 homo sapien	941	36	24.5	1622	5	Q06550	Q06550 cryptospori
869	36	24.5	577	10	O64546	O64546 arabidopsis	942	36	24.5	1637	6	Q9XSV8	Q9XSV8 bos taurus
870	36	24.5	579	11	Q9WV10	Q9WV10 mus musculus	943	36	24.5	1661	4	O9Y4G4	O9Y4G4 homo sapien
871	36	24.5	582	12	Q98683	Q98683 simian cyto	944	36	24.5	1715	5	Q9XU09	Q9XU09 bombyx mori
872	36	24.5	584	4	Q9Y573	Q9Y573 homo sapien	945	36	24.5	1783	11	O70395	O70395 mus musculus
873	36	24.5	585	11	P70581	P70581 rattus norv	946	36	24.5	1872	11	P70208	P70208 mus musculus
874	36	24.5	596	4	O00185	O00185 homo sapien	947	36	24.5	1873	12	Q83044	Q83044 lettuce inf
875	36	24.5	602	11	O63485	O63485 rattus norv	948	36	24.5	1907	4	O15017	O15017 homo sapien
876	36	24.5	604	2	O53124	O53124 mycobacteri	949	36	24.5	2140	3	Q02316	Q02316 lentinula e
877	36	24.5	615	2	Q92426	Q92426 streptomyce	950	36	24.5	2163	5	O01912	O01912 caenorhabdi
878	36	24.5	615	10	Q9Z5T2	Q9Z5T2 citrus para	951	36	24.5	2167	5	O76840	O76840 caenorhabdi
879	36	24.5	624	4	Q9Y5P4	Q9Y5P4 homo sapien	952	36	24.5	2185	3	Q12721	Q12721 ustilago ma
880	36	24.5	625	12	O83890	O83890 ovine adeno	953	36	24.5	2215	5	Q22465	Q22465 caenorhabdi
881	36	24.5	640	11	O09182	O09182 rattus norv	954	36	24.5	2218	12	O09705	O09705 lassa virus
882	36	24.5	646	5	Q23521	Q23521 caenorhabdi	955	36	24.5	2321	4	O9Y6L8	O9Y6L8 homo sapien
883	36	24.5	649	2	Q92D10	Q92D10 rickettsia	956	36	24.5	2447	4	Q22463	Q22463 caenorhabdi
884	36	24.5	649	10	O81620	O81620 arabidopsis	957	36	24.5	2482	6	Q28833	Q28833 sus scrofa
885	36	24.5	688	8	Q34312	Q34312 dictyosteli	958	36	24.5	3085	4	O00183	O00183 homo sapien
886	36	24.5	699	4	O99876	O99876 homo sapien	959	36	24.5	3418	4	O13879	O13879 homo sapien
887	36	24.5	710	10	Q9XED1	Q9XED1 arabidopsis	960	36	24.5	3456	12	P89201	P89201 sweet potat
888	36	24.5	713	2	O52672	O52672 escherichia	961	36	24.5	3623	4	O60494	O60494 homo sapien
889	36	24.5	717	5	O9Y169	O9Y169 drosophila	962	36	24.5	4292	4	Q15141	Q15141 homo sapien
890	36	24.5	722	3	Q9Y7V4	Q9Y7V4 sclerotinia	963	36	24.5	4302	4	Q15140	Q15140 homo sapien
891	36	24.5	733	3	O24037	O24037 drosophila	964	36	24.5	4472	2	O33954	O33954 streptomyce
892	36	24.5	747	5	Q20046	Q20046 caenorhabdi	965	36	24.5	6658	5	O76281	O76281 drosophila
893	36	24.5	752	4	O75799	O75799 homo sapien	966	36	24.5	7962	4	O10465	O10465 homo sapien
894	36	24.5	759	2	Q9X4P3	Q9X4P3 pseudomonas	967	35.5	24.1	62	2	Q47407	Q47407 escherichia
895	36	24.5	782	5	O44500	O44500 caenorhabdi	968	35.5	24.1	66	2	O69083	O69083 pseudomonas


```

969 35.5 24.1 111 1 Q9YEH4
970 35.5 24.1 136 2 P73378
971 35.5 24.1 170 1 O08448
972 35.5 24.1 184 10 Q9XE78
973 35.5 24.1 202 10 Q81842
974 35.5 24.1 225 10 Q41800
975 35.5 24.1 227 2 Q9ZGR2
976 35.5 24.1 237 5 O01599
977 35.5 24.1 240 10 Q39088
978 35.5 24.1 262 13 Q98988
979 35.5 24.1 271 3 O13381
980 35.5 24.1 285 5 O17997
981 35.5 24.1 304 10 O80864
982 35.5 24.1 310 2 Q9Z9S6
983 35.5 24.1 315 2 Q9Z9T1
984 35.5 24.1 315 10 Q9XGC3
985 35.5 24.1 316 5 O16463
986 35.5 24.1 328 10 Q24463
987 35.5 24.1 331 4 Q75480
988 35.5 24.1 332 5 Q93868
989 35.5 24.1 335 2 Q9X4M6
990 35.5 24.1 342 10 Q9X1Z6
991 35.5 24.1 343 10 Q49336
992 35.5 24.1 347 4 O60619
993 35.5 24.1 371 13 Q9W676
994 35.5 24.1 372 3 Q9Y8H2
995 35.5 24.1 373 4 Q43679
996 35.5 24.1 375 5 Q20002
997 35.5 24.1 379 11 Q54897
998 35.5 24.1 380 10 Q39540
999 35.5 24.1 385 13 Q73713
1000 35.5 24.1 388 12 Q41927

```

ALIGNMENTS

```

RESULT 1
Q7752 ID 077752 PRELIMINARY; PRT; 120 AA.
AC 07752;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE GONADOTROPHIN ALPHA SUBUNIT.
OS Trichosurus vulpecula (Brush-tailed possum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PIUITARY;
RX MEDLINE; 98349946.
RA FIDLER A.E., LAWRENCE S.B., VANMONTFORT D.M., TISDALL D.J.,
RA MCNATTY K.P.;
RT "The Australian brushtail possum (Trichosurus vulpecula) gonadotrophin
alpha-subunit: analysis of cDNA sequence and pattern of expression.";
RL J. Mol. Endocrinol. 20:345-353(1998).
DR EMBL; AF004520; AAC63900.1; -.
DR HSSP; P01215; 1HRP.
DR PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
DR PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
DR PFAM; PF00236; hormone6; 1.
DR PRINTS; PR00274; GLYCOHORMONE.
DR PRINTS; PR00438; GFCYSKNOT.
SQ SEQUENCE 120 AA; 13499 MW; C1BAADCF CRC32;

```

```

Query Match 56.5%; Score 83; DB 6; Length 120;
Best Local Similarity 57.1%; Pred. No. 4e-05;
Matches 16; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

```

```

QY 2 pd----vqdcpectlqenpfsgqgapi 25
|| :| |||| :|| :|| :||

```

```

Db 26 PDGEFTMQCPECKLKNKYFSKVGAPI 53
RESULT 2
Q9XSW8 ID 09XSW8 PRELIMINARY; PRT; 120 AA.
AC 09XSW8;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE THYROTROPIN ALPHA SUBUNIT.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PIUITARY;
RA YANG X., MCGRAW R.A., FERGUSON D.C.;
RT "cDNA cloning of canine thyrotropin alpha subunit gene.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF160250; RAD42900.1; -.
DR PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
DR PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
DR PRINTS; PR00274; GLYCOHORMONE.
SQ SEQUENCE 120 AA; 13411 MW; 132E5DC6 CRC32;

```

```

Query Match 55.8%; Score 82; DB 6; Length 120;
Best Local Similarity 57.1%; Pred. No. 5.6e-05;
Matches 16; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

```

```

QY 2 pd----vqdcpectlqenpfsgqgapi 25
|| :| |||| :|| :|| :||
Db 26 PDGEFTMQCPECKLKNKYFSKLGAPI 53

```

```

RESULT 3
Q46642 ID 046642 PRELIMINARY; PRT; 120 AA.
AC 046642;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE LUTEINIZING HORMONE ALPHA SUBUNIT PRECURSOR.
GN LH.
OS Equus burchelli (Plains zebra) (Equus quagga).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Perissodactyla; Equidae; Equus.
RN [1]
RP SEQUENCE FROM N.A.
RA CHOPINEAU M., MARTINAT N., POURCHET C., GUILLOU F.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y16326; CAA76177.1; -.
DR HSSP; P01215; 1HRP.
DR PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
DR PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
DR PFAM; PF00236; hormone6; 1.
DR PRINTS; PR00274; GLYCOHORMONE.
KW Signal.
FT SIGNAL.
FT CHAIN 1 24 POTENTIAL.
FT CHAIN 25 120 LUTEINIZING HORMONE ALPHA SUBUNIT.
SQ SEQUENCE 120 AA; 13720 MW; 5DFE37CA CRC32;

```

```

Query Match 52.4%; Score 77; DB 6; Length 120;
Best Local Similarity 53.6%; Pred. No. 0.0003;
Matches 15; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

```

```

QY 2 pd----vqdcpectlqenpfsgqgapi 25
|| :| |||| :|| :|| :||
Db 26 PDGEFTQDCPECKLKNKYFSKLGAPI 53

```

```

RESULT 4

```



```

Q90286 ID Q90286 PRELIMINARY; PRT; 107 AA.
AC Q90286;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE GONADOTROPIN ALPHA SUBUNIT (FRAGMENT).
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprininae; Cyprininae; Carassius.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97242867.
RA KOBAYASHI M., KATO Y., YOSHURA Y., AIDA K.;
RT "Molecular cloning of cDNA encoding two types of pituitary
RT gonadotropin alpha subunit from the goldfish, Carassius auratus.";
RL Gen. Comp. Endocrinol. 105:372-378(1997).
DR EMBL; D86551; BAA1311.1; -.
DR HSP; P01215; IHRP.
DR PFAM; PF00236; hormone6; 1.
FT NON_TER 1
FT NON_TER 107
SQ SEQUENCE 107 AA; 11959 MW; 880C6840 CRC32;

Query Match 49.7%; Score 73; DB 13; Length 107;
Best Local Similarity 63.2%; Pred. No. 0.001;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 7 cpecltqlenpfssqgapi 25
Db 28 CEECKLKENNIFSKPGAPV 46

RESULT 5
Q90287 ID Q90287 PRELIMINARY; PRT; 108 AA.
AC Q90287;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE GONADOTROPIN ALPHA SUBUNIT (FRAGMENT).
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprininae; Cyprininae; Carassius.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97242867.
RA KOBAYASHI M., KATO Y., YOSHURA Y., AIDA K.;
RT "Molecular cloning of cDNA encoding two types of pituitary
RT gonadotropin alpha subunit from the goldfish, Carassius auratus.";
RL Gen. Comp. Endocrinol. 105:372-378(1997).
DR EMBL; D86552; BAA1311.1; -.
DR HSP; P01215; IHRP.
DR PFAM; PF00236; hormone6; 1.
FT NON_TER 1
FT NON_TER 108
SQ SEQUENCE 108 AA; 12129 MW; B2B1A212 CRC32;

Query Match 49.7%; Score 73; DB 13; Length 108;
Best Local Similarity 63.2%; Pred. No. 0.001;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 7 cpecltqlenpfssqgapi 25
Db 29 CEECKLKENNIFSKPGAPV 47

RESULT 6
Q9YGP3 ID Q9YGP3 PRELIMINARY; PRT; 116 AA.
AC Q9YGP3;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE GONADOTROPIN ALPHA SUBUNIT GLYCOPROTEIN PRECURSOR.
OS Ictalurus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Siluriformes;
OC Ictaluridae; Ictalurus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-KANSAS;
RX MEDLINE; 97430286.
RA LIU Z., LI P., ARGUE B.J., DUNHAM R.A.;
RT "Gonadotropin alpha-subunit glycoprotein from channel catfish
RT (Ictalurus punctatus) and its expression during hormone-induced
RT ovulation.";
RL Mol. Mar. Biol. Biotechnol. 6:217-227(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-KANSAS;
RX LIU Z.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF112190; AAD18004.1; -.
DR HSP; P01215; IHRP.
DR PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
DR PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
SQ SEQUENCE 116 AA; 13089 MW; 1BAB9CA7 CRC32;

Query Match 49.7%; Score 73; DB 13; Length 116;
Best Local Similarity 63.2%; Pred. No. 0.0011;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 7 cpecltqlenpfssqgapi 25
Db 32 CEECKLKENNIFSKPGAPV 50

RESULT 7
Q91370 ID Q91370 PRELIMINARY; PRT; 119 AA.
AC Q91370;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE GONADOTROPIN ALPHA 1 SUBUNIT.
OS Oncorhynchus masou (Cherry salmon) (Masu salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
OC Salmoniformes; Salmonidae; Oncorhynchus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94197892.
RA GEN K., MARDYAMA O., KATO T., TOMIZAWA K., WAKABAYASHI K., KATO Y.;
RT "Molecular cloning of cDNAs encoding two types of gonadotropin alpha
RT subunit from the masu salmon, Oncorhynchus masou: construction of
RT specific oligonucleotides for the alpha 1 and alpha 2 subunits.";
RL J. Mol. Endocrinol. 11:265-273(1993).
DR EMBL; S69273; AAB30421.1; -.
DR HSP; P01215; IHRP.
DR PROSITE; PS00779; GLYCO_HORMONE_ALPHA_1; 1.
DR PROSITE; PS00780; GLYCO_HORMONE_ALPHA_2; 1.
DR PFAM; PF00236; hormone6; 1.
DR PRINTS; PR00274; GLYCOHORMONE.
SQ SEQUENCE 119 AA; 13131 MW; 2EEF28F4 CRC32;

Query Match 49.7%; Score 73; DB 13; Length 119;
Best Local Similarity 63.2%; Pred. No. 0.0012;
Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```



```

Qy 2 pdvqdcpectlqenpf 17
   |||:|:|:|
Db 36 PDTRFCQCGLGKNAF 51

RESULT 16
Q9XWZ5
ID Q9XWZ5 PRELIMINARY; PRT; 114 AA.
AC Q9XWZ5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE Y42A5A.3 PROTEIN.
GN Y42A5A.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA STWARD C.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFFEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL; AL032618; CAA21486.1; -. EBBICIA77 CRC32;
SQ SEQUENCE 114 AA; 12879 MW; 12879 MW; 12879 MW; 12879 MW;

Query Match 32.0%; Score 47; DB 5; Length 114;
Best Local Similarity 50.0%; Pred. No. 7.1;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 8 pectlqenpfsgqapi 25
   |||:|:|:|
Db 65 PIVTRQENTFIDODGVPV 82

RESULT 17
Q27481
ID Q27481 PRELIMINARY; PRT; 1113 AA.
AC Q27481;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE C47E12.5 PROTEIN.
GN C47E12.5
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
[1]
RN SEQUENCE FROM N.A.
RX COLES L.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE; 94150718.

```

```

RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFFEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL; Z68882; CAA93101.1; -.
DR PFAM; PF00899; Thif_family; 2.
SQ SEQUENCE 1113 AA; 124083 MW; 2CA44E4B CRC32;

Query Match 32.0%; Score 47; DB 5; Length 1113;
Best Local Similarity 29.7%; Pred. No. 66;
Matches 11; Conservative 4; Mismatches 6; Indels 16; Gaps 1;

Qy 2 pdvqdcpectlqenp-----ffsqpg 22
   |||:|:|:|
Db 673 PPEKEIPVCTLNKFPNEIQTIOWAREQFETFEAQP 709

RESULT 18
Q19919
ID Q19919 PRELIMINARY; PRT; 242 AA.
AC Q19919;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-1998 (TrEMBLrel. 06, Last annotation update)
DE HYPOTHEICAL 26.2 KD PROTEIN F31A3.1 IN CHROMOSOME X.
GN F31A3.1
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
[1]
RN SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2;
RA MURRAY J., LE T.T.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO FAMILY UPF.
DR EMBL; U58742; AAB36856.1; -.
DR WORMPEP; F31A3.1; CE07158.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 3 23 POTENTIAL.
SQ SEQUENCE 242 AA; 26213 MW; 0B4E01F0 CRC32;

Query Match 31.6%; Score 46.5; DB 5; Length 242;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 12; Conservative 1; Mismatches 10; Indels 1; Gaps 1;

Qy 1 apdvqdcpectlqenpfsgqap 24
   |||:|:|:|
Db 60 AP-VQHVPQCQCQCAPOCQCPAAP 82

RESULT 19
O66902
ID O66902 PRELIMINARY; PRT; 746 AA.
AC O66902;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE TRANSCRIPTIONAL REGULATORY PROTEIN HYPF.
GN HYPF.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
[1]

```


RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE: 98196666.
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus";
 RL Nature 392:353-358(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF000701; AAC06863.1; -
 DR HSSP: P00818; IAPS.
 DR PFAM: PF00708; Acylphosphatase; 1.
 DR PFAM: PF01300; Sua5_YC10_YrdC; 1.
 SQ SEQUENCE 746 AA; 84761 MW; AEELCBDF CRC32;

Query Match 31.6%; Score 46.5; DB 2; Length 746;
 Best Local Similarity 50.0%; Pred. No. 53;
 Matches 11; Conservative 2; Mismatches 4; Indels 5; Gaps 2;

QY 7 cpecltq-enp-----ffsqgpa 23
 |||| : ||| : ||| :
 Db 159 CPECKREYENPLDRRHFAPNA 180

RESULT 20
 O22474 PRELIMINARY; PRT; 227 AA.
 AC O22474;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
 DE S27-RNASE.
 GN S.
 OS Malus domestica (Apple) (Malus sylvestris).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Rosidae; eurosids I; Rosales; Rosaceae; Malus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA VERDOODT L., VAN HAUTE A., GODERIS I.J., DE WITTE K., KEULEMANS J.,
 RA BROTHAERTS W.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF016919; AAB70516.1; -
 DR MENDEL; 26367; Maldo; 1284; 26367.
 DR PFAM: PF00445; ribonuclease T2; 1.
 SQ SEQUENCE 227 AA; 25991 MW; 1B9BD8A5 CRC32;

Query Match 31.3%; Score 46; DB 10; Length 227;
 Best Local Similarity 52.4%; Pred. No. 20;
 Matches 11; Conservative 1; Mismatches 1; Indels 8; Gaps 2;

QY 6 dpecltqenpfsgpapll 26
 ||| ||| ||| : ||| :
 Db 207 DCP-----NPF--QPGSPYL 219

RESULT 21
 Q9Y3I5 PRELIMINARY; PRT; 433 AA.
 AC Q9Y3I5;
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
 DE DA191P20.2 (NOVEL FIBRONECTIN TYPE III DOMAIN CONTAINING PROTEIN

DE SIMILAR TO RING FINGER PROTEIN MID1 (MIDLINE 1)) (FRAGMENT).
 GN DA191P20.2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN [1]
 RA WILSON S.;
 RP SEQUENCE FROM N.A.
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL034399; CAB41046.1; -
 FT NON_TER 1
 SQ SEQUENCE 433 AA; 49294 MW; 9CF24DCF CRC32;

Query Match 31.0%; Score 45.5; DB 4; Length 433;
 Best Local Similarity 45.8%; Pred. No. 44;
 Matches 11; Conservative 2; Mismatches 6; Indels 5; Gaps 1;

QY 1 apdvqdcpe-----ctlgpenffs 19
 |||| ||| ||| : ||| : ||| :
 Db 404 APDFIDYPERQECNCRPQSPYVS 427

RESULT 22
 Q65553 PRELIMINARY; PRT; 3247 AA.
 AC Q65553;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
 DE UL36.
 GN UL36.
 OS Bovine herpesvirus 1.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COOPER;
 RA SCHWYZER M., VLCEK C., LOWERY D.E., BELLO L.J., MEYER G., MISRA V.,
 RA THIRY E., PACES V.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-179 FROM N.A.
 RC STRAIN=COOPER;
 RA SCHWYZER M., STYGER D., VOGT B., LOWERY D.E., SIMARD C.,
 RA LABOISSIERE S., MISRA V., VLCEK C., PACES V.;
 RL Vet. Microbiol. 0:0-0(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COOPER;
 RA SCHWYZER M.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-179 FROM N.A.
 RC STRAIN=JURA;
 RX MEDLINE: 97164286.
 RA SCHWYZER M., STYGER D., VOGT B., LOWERY D.E., SIMARD C.,
 RA LABOISSIERE S., MISRA V., VLCEK C., PACES V.;
 RT "Gene contents in a 31-kb segment at the left genome end of bovine
 RT herpesvirus-1";
 RN Vet. Microbiol. 53:67-77(1996).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JURA;
 RA SCHWYZER M., VLCEK C., LOWERY D.E., BELLO L.J., MEYER G., MISRA V.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JURA;
 RA SCHWYZER M.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: 278205; CAB01605.1; -
 DR EMBL: AJ004801; CAA06097.1; -

RESULT 28
O14873
ID O14873
PRELIMINARY:
PRT: 228 AA.

RP MUTAGENESIS.
RX MEDLINE; 98153115.

[illegible]

OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
 OC Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae;
 OC Xenopus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE; 95369846.
 RA KATO Y., SALTER-CID L., FLAJNIK M., NAMIKAWA C., SASAKI M., NONAKA M.;
 RT "Duplication of the MHC-linked Xenopus complement factor B gene."
 RL Immunogenetics 42:196-203(1995).
 DR EMBL; D49373; BAA08371.1; -
 DR HSSP; P00734; IUVT.
 DR PFAM; PF00084; sushi; 2.
 DR PFAM; PF00089; trypsin; 1.
 DR PFAM; PF00092; vwa; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 SQ SEQUENCE 745 AA; 84313 MW; F6A0FAD0 CRC32;

Query Match 29.9%; Score 44; DB 13; Length 745;
 Best Local Similarity 44.4%; Pred. No. 1.2e+02;
 Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

OY 7 cpecltqenpfssqgap 24
 ||: ||: ||||
 DB 573 CLSCTMGTAQVLKQPGAP 590

RESULT 41
 O14830 PRELIMINARY; PRT; 753 AA.
 AC O14830;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE PROTEIN PHOSPHATASE WITH EF-HANDS-2 LONG FORM.
 GN PPEF-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=RETINA;
 RX MEDLINE; 97471020.
 RA SHERMAN P.M., SUN H., MACKE J.P., WILLIAMS J., SMALLWOOD P.M.,
 RA NATHANS J.;
 RT "Identification and characterization of a conserved family of protein
 serine/threonine phosphatases homologous to Drosophila retinal
 degeneration C."
 RL Proc. Natl. Acad. Sci. U.S.A. 94:11639-11644(1997).
 DR EMBL; AF023456; AAB82796.1; -
 DR HSSP; P08129; 1FJM.
 DR PFAM; PF00036; ehand; 2.
 DR PFAM; PF00612; IQ; 1.
 DR PFAM; PF00149; Stphphatase; 2.
 DR PRINTS; PR00114; STPHPHATASE.
 SQ SEQUENCE 753 AA; 86430 MW; 58FF1915 CRC32;

Query Match 29.9%; Score 44; DB 4; Length 753;
 Best Local Similarity 38.1%; Pred. No. 1.2e+02;
 Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 3 dvqdcpcltqenpfssqga 23
 |::|::|::|::|::|
 DB 732 DASECPOATNAKDCGCSPPGA 752

RESULT 42
 Q61161 PRELIMINARY; PRT; 821 AA.
 ID Q61161
 AC Q61161;

DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE RAB8 INTERACTING PROTEIN.
 GN RAB8IP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C;
 RX MEDLINE; 96209873.
 RA REN M., ZENG J., DELEMOS-CHIARANDINI C., ROSENFELD M., ADESNIK M.,
 RA SABATINI D.D.;
 RT "In its active form, the GTP-binding protein rab8 interacts with a
 stress-activated protein kinase."
 RL Proc. Natl. Acad. Sci. U.S.A. 93:5151-5155(1996).
 DR EMBL; U50595; AAC52571.1; -
 DR MGD; MGI:108033; Rab8ip.
 DR PFAM; PF00780; CNH; 1.
 DR PFAM; PF00069; pkinase; 1.
 DR PFAM; PF00069; pkinase; 1.
 SQ SEQUENCE 821 AA; 91265 MW; 2833A9E2 CRC32;

Query Match 29.9%; Score 44; DB 11; Length 821;
 Best Local Similarity 37.5%; Pred. No. 1.4e+02;
 Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 2 pdvqdcpcltqenpf 17
 ||: ||: ||:
 DB 594 PDTKGLQCRVVRNPY 609

RESULT 43
 O01768 PRELIMINARY; PRT; 2180 AA.
 ID O01768
 AC O01768;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE SIMILARITY TO EGF-LIKE DOMAINS.
 GN T2IE3.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE; 94150718.
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., FULTON L.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., JER M., JOHNSTON L.,
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA SKALSON J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SWALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans."
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX DU Z., LE T.T.;
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA WATERSTON R.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF003133; AAB54138.1; -


```
DR HSP; Q07954; ICR8.
DR PROSITE; PS01209; LDLRA_1; 10.
DR PFAM; PF00008; EGF; 9.
DR PFAM; PF00057; ldl_recept_a; 18.
DR PFAM; PF00058; ldl_recept_b; 8.
DR PRINTS; PR00261; LDLRECEPTOR.
KW Glycoprotein.
SQ SEQUENCE 2180 AA; 241705 MW; 112867E3 CRC32;

Query Match 29.9%; Score 44; DB 5; Length 2180;
Best Local Similarity 39.3%; Pred. NO. 3.5e+02;
Matches 11; Conservative 1; Mismatches 10; Indels 6; Gaps 1;

QY 4 vqdc-----pectlqenpfssqgapi 25
   :|| || || || || || || || || ||
Db 37 INDCANFADESNCTCNEENFCQSGACI 64

RESULT 44
O61291 PRELIMINARY; PRT; 4545 AA.
ID Q61291
AC Q61291;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE LOW DENSITY LIPOPROTEIN RECEPTOR RELATED PROTEIN (AM2 RECEPTOR).
GN LRP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 93250049.
RA VAN LEEUWEN F.;
RT "Molecular cloning and sequencing of the murine alpha-2-macroglobulin
  receptor cDNA.";
RL Biochim. Biophys. Acta 1173:71-74(1993).
DR EMBL; X67469; CAA47817.1; -.
DR HSP; Q07954; ICR8.
DR MGD; MGI:96828; Lrp.
DR PROSITE; PS01187; EGF_CA; 2.
DR PROSITE; PS01209; LDLRA_1; 26.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PFAM; PF00008; EGF; 16.
DR PFAM; PF00057; ldl_recept_a; 31.
DR PFAM; PF00058; ldl_recept_b; 34.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00261; LDLRECEPTOR.
KW Glycoprotein; EGF-like domain.
SQ SEQUENCE 4545 AA; 504761 MW; 9D57A4D4 CRC32;

Query Match 29.9%; Score 44; DB 11; Length 4545;
Best Local Similarity 53.3%; Pred. NO. 7.2e+02;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 dvqdcpectlqenpf 17
   :||| || || ||
Db 3445 DERDCPEVTCAPNQF 3459

RESULT 45
O57381 PRELIMINARY; PRT; 735 AA.
ID O57381
AC O57381;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE BONE MORPHOGENETIC PROTEIN 1B.
GN BMP-1.
OS Xenopus laevis (African clawed frog).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae;
OC Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RA GOODMAN S., ALBANO R., MATTHEWS G., TANNAHILL D., DALE L.;
RL Dev. Biol. 0:0-0(0).
DR EMBL; I09660; CAA70853.1; -.
DR HSP; P00736; LAPO.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PFAM; PF01400; Astacin; 1.
DR PFAM; PF00008; EGF; 1.
DR PFAM; PF00431; CUB; 3.
DR PRINTS; PR00480; ASTACIN.
KW Glycoprotein; EGF-like domain.
SQ SEQUENCE 735 AA; 83575 MW; 33E4E532 CRC32;
```

```
Query Match 29.6%; Score 43.5; DB 13; Length 735;
Best Local Similarity 52.4%; Pred. NO. 1.4e+02;
Matches 11; Conservative 2; Mismatches 5; Indels 3; Gaps 2;

QY 7 cpec--tlqenp-ffsqpgap 24
   || | |||: || || |
Db 309 CPACGETLQDSQGNFSSPGFP 329
```

```
RESULT 46
O91925 PRELIMINARY; PRT; 977 AA.
ID O91925
AC O91925;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE XTLD PROTEIN.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae;
OC Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRYO;
RA LIN J., MAEDA R., ONG R., KIM J., KUNG H., MAENO M.;
RT "xtld, a Xenopus homolog of dorso-ventral polarity gene in Drosophila,
  modifies tissue phenotypes of the ventral mesoderm.";
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D83476; BAAL1922.1; -.
DR HSP; P00736; LAPO.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS01187; EGF_CA; 2.
DR PFAM; PF01400; Astacin; 1.
DR PFAM; PF00008; EGF; 2.
DR PFAM; PF00431; CUB; 5.
DR PRINTS; PR00480; ASTACIN.
KW Glycoprotein; EGF-like domain.
SQ SEQUENCE 977 AA; 110199 MW; 868BA492 CRC32;
```

```
Query Match 29.6%; Score 43.5; DB 13; Length 977;
Best Local Similarity 52.4%; Pred. NO. 1.9e+02;
Matches 11; Conservative 2; Mismatches 5; Indels 3; Gaps 2;

QY 7 cpec--tlqenp-ffsqpgap 24
   || | |||: || || |
Db 310 CPACGETLQDSQGNFSSPGFP 330
```

```
RESULT 47
O57382 PRELIMINARY; PRT; 1019 AA.
ID O57382
AC O57382;
```


DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE XOLLOID.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
 OC Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae;
 OC Xenopus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-OVERLY;
 RA GOODMAN S., ALBANO R., MATTHEWS G., TANNAHILL D., DALE L.;
 RL Dev. Biol. 0:0-0(0).
 DR EMBL: Y09661; CAA70854.1; -.
 DR HSP; P00736; IAPQ.
 DR PROSITE: PS00010; ASX_HYDROXYL; 2.
 DR PROSITE: PS01187; EGF_CA; 2.
 DR PFAM: PF01400; Astacin; 1.
 DR PFAM: PF00008; EGF; 2.
 DR PFAM: PF00431; CUB; 5.
 DR PRINTS: PR00480; ASTACIN.
 KW Glycoprotein; EGF-like domain.
 SQ SEQUENCE 1019 AA; 114891 MW; 284667A2 CRC32;

Query Match 29.6%; Score 43.5; DB 13; Length 1019;

Best Local Similarity 44.4%; Pred. No. 2e+02; Length 1019;
 Matches 12; Conservative 3; Mismatches 9; Indels 3; Gaps 2;

QY 1 apdvqdcpcet--tlgenp-ffsqpgap 24

Db 345 AKLYKCPACGETLQDSSGNFSAPGYP 371

RESULT 48

ID Q47891 PRELIMINARY; PRT; 272 AA.
 AC Q47891;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE HYPOTHETICAL 31.9 KD PROTEIN.
 OS Fremyella diplosiphon (Calothrix PCC 7601).
 OC Bacteria; Cyanobacteria; Nostocales; Rivulariaceae; Fremyella.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TRANSPOSON-TN5469;
 RX MEDLINE; 96099283.
 RA KAHN K., SCHAEFER M.R.;
 RT "Characterization of transposon Tn5469 from the cyanobacterium Fremyella diplosiphon."
 RL J. Bacteriol. 177:7026-7032(1995).
 DR EMBL: U33002; AAA92004.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 272 AA; 31923 MW; 95E6DB02 CRC32;

Query Match 29.3%; Score 43; DB 2; Length 272;

Best Local Similarity 34.8%; Pred. No. 64; Length 272;
 Matches 8; Conservative 6; Mismatches 7; Indels 2; Gaps 1;

QY 2 pdvqdcpcetlqenpf--fsqpg 22

Db 178 PLIECPHCHEKFPPLSRYSRFG 200

RESULT 49

ID O70233 PRELIMINARY; PRT; 307 AA.
 AC O70233;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)

DE PHOSPHOLIPID SCRAMBLASE 1 (PL SCRAMBLASE) (CA2+ DEPENDENT PHOSPHOLIPID SCRAMBLASE).
 GN PLSCL1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-FIBROBLAST;
 RX MEDLINE; 98153115.
 RA ZHOU Q., SIMS P.J., WIEDMER T.;
 RL "Identity of a conserved motif in phospholipid scramblase that is required for Ca2+-accelerated transbilayer movement of membrane phospholipids".
 RL Biochemistry 37:2356-2360(1998).
 CC -!- FUNCTION: THIS PLASMA MEMBRANE PROTEIN MEDIATES ACCELERATED TRANSBILAYER MIGRATION OF PHOSPHOLIPIDS UPON BINDING CALCIUM IONS.
 CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (POTENTIAL).
 DR EMBL: AF015790; AAC40053.1; -.
 DR MGD; MGI:1270860; PLSCL1.
 KW Apoptosis; Transmembrane; Phosphorylation; Calcium-binding.
 FT MOD_RES 159 PHOSPHORYLATION (BY PKC) (POTENTIAL).
 FT CA_BIND 271 POTENTIAL.
 SQ SEQUENCE 307 AA; 34080 MW; 55475815 CRC32;

Query Match 29.3%; Score 43; DB 11; Length 307;

Best Local Similarity 47.4%; Pred. No. 72; Length 307;
 Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 7 cpecltqenpfssqpgapi 25

Db 183 CPCCQLQEIEIQAPGVPI 201

RESULT 50

ID Q23852 PRELIMINARY; PRT; 421 AA.
 AC Q23852;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE HYPOTHETICAL PROTEIN (FRAGMENT).
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Dictyostelida; Dictyostelium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AX3;
 RX MEDLINE; 91002566.
 RA GIORDA R., OHMACHI T., SHAW D.R., ENNIS H.L.;
 RT "A shared internal threonine-glutamic acid-threonine-proline repeat defines a family of Dictyostelium discoideum spore germination specific proteins".
 RT Biochemistry 29:7264-7269(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AX3;
 RA GIORDA R., OHMACHI T., SHAW D.R., ENNIS H.L.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U20661; AAB34076.1; -.
 DR PFAM: PF00069; pkinase; 1.
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 421 AA; 47860 MW; 16DE3A1C CRC32;

Query Match 29.3%; Score 43; DB 5; Length 421;

Best Local Similarity 45.0%; Pred. No. 98; Length 421;
 Matches 9; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

QY 2 pdvqdcpcetlqenpfssqpg 21

Db 262 PPVETPKPTI--NPFNTPT 279

Search completed: July 14, 2000, 09:35:54
Job time: 2405 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 14, 2000, 08:55:43 ; Search time 37.7 Seconds
(without alignments)
18.848 Million cell updates/sec

Title: BETA-CHAIN

Perfect score: 171

Sequence: 1 vctyrdfiyrtveipgclhvpafsyypva 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters: 188960

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 99%

Listing first 1000 summaries

Database : A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match %	Length	DB ID	Description
1	163	95.3	112	1	R31002	Human thyroid stim
2	163	95.3	132	1	W99549	hTSH-beta analogue
3	163	95.3	132	1	W99550	hTSH-beta analogue
4	156.5	91.5	137	1	R39419	TSH beta subunit.
5	131	76.6	145	1	R15092	hCG/hTSH chimera,
6	121	70.8	145	1	R15091	hCG/hTSH chimera,
7	104	60.8	145	1	W27683	Chorionic gonadotr
8	99	57.9	142	1	P94151	Fish gonadotropin
9	96	56.1	145	1	W27686	Chorionic gonadotr
10	95	55.6	141	1	P60602	Sequence of human
11	95	55.6	141	1	W99548	hLH-beta analogue
12	95	55.6	141	1	W99547	hLH-beta analogue
13	95	55.6	141	1	W99520	Glycoprotein hormo
14	95	55.6	141	1	W99521	Glycoprotein hormo
15	95	55.6	141	1	W99522	Glycoprotein hormo
16	95	55.6	141	1	W99523	Glycoprotein hormo
17	95	55.6	141	1	W99524	Glycoprotein hormo
18	95	55.6	141	1	W99525	Glycoprotein hormo
19	95	55.6	141	1	W99516	Glycoprotein hormo
20	95	55.6	141	1	W99517	Glycoprotein hormo
21	95	55.6	141	1	W99519	Glycoprotein hormo
22	95	55.6	234	1	R86261	Partially deglycos
23	95	55.6	234	1	R86249	Single chain gonad
24	92	53.8	129	1	R10038	Ovine FSH beta sub
25	92	53.8	145	1	R15066	hCG/hFSH chimera,
26	92	53.8	242	1	P61785	Sequence of porcine
27	91	53.2	76	1	R68979	Human FSH beta sub
28	91	53.2	76	1	R73939	Human FSH-beta exo
29	91	53.2	110	1	P92145	Beta subunit of ur
30	91	53.2	111	1	R31000	Human follicle sti
31	91	53.2	112	1	R15086	hCG/hFSH chimera,
32	91	53.2	118	1	R15082	hCG/hFSH chimera,
33	91	53.2	129	1	R06724	Human follicle-sti

34	91	53.2	129	1	R10096	Engineered partial
35	91	53.2	129	1	P60618	Sequence of mature
36	91	53.2	129	1	R30786	Follicle stimulat
37	91	53.2	129	1	R99418	FSH beta subunit.
38	91	53.2	129	1	W19607	Human beta follicl
39	91	53.2	129	1	W19123	Human beta follicl
40	91	53.2	129	1	W52250	Human FSH-beta pro
41	91	53.2	129	1	W62027	Human beta follicl
42	91	53.2	129	1	W89411	Human follicle sti
43	91	53.2	129	1	W99543	hFSH-beta analogue
44	91	53.2	129	1	W99546	hFSH-beta analogue
45	91	53.2	129	1	W99505	Glycoprotein hormo
46	91	53.2	129	1	W99499	Glycoprotein hormo
47	91	53.2	129	1	W99500	Glycoprotein hormo
48	91	53.2	129	1	W99501	Glycoprotein hormo
49	91	53.2	129	1	W99502	Glycoprotein hormo
50	91	53.2	129	1	W99503	Glycoprotein hormo
51	91	53.2	129	1	W99504	Glycoprotein hormo
52	91	53.2	129	1	W99496	Glycoprotein hormo
53	91	53.2	129	1	W99497	Glycoprotein hormo
54	91	53.2	129	1	W99498	Glycoprotein hormo
55	91	53.2	137	1	W99527	Glycoprotein hormo
56	91	53.2	139	1	R15085	hCG/hFSH chimera,
57	91	53.2	141	1	W99518	Glycoprotein hormo
58	91	53.2	145	1	R15115	hCG/hLH chimera, A
59	91	53.2	145	1	W27684	Chorionic gonadotr
60	91	53.2	147	1	R15090	hCG/hFSH chimera,
61	91	53.2	222	1	R86268	Partially deglycos
62	91	53.2	222	1	R86256	Single chain gonad
63	91	53.2	226	1	R86267	Partially deglycos
64	91	53.2	226	1	R86255	Single chain gonad
65	91	53.2	229	1	R86262	Partially deglycos
66	91	53.2	229	1	R86250	Single chain gonad
67	90	52.6	122	1	R15123	hCG/hLH chimera, A
68	90	52.6	131	1	W47027	eCG hormone beta-s
69	90	52.6	134	1	W47025	eCG hormone beta-s
70	90	52.6	137	1	W47026	eCG hormone beta-s
71	90	52.6	139	1	W47024	eCG hormone beta-s
72	90	52.6	145	1	R15116	hCG/hLH chimera, A
73	90	52.6	145	1	R15117	hCG/hLH chimera, A
74	90	52.6	145	1	R15122	hCG/hLH chimera, A
75	90	52.6	145	1	R15124	hCG/hLH chimera, A
76	90	52.6	169	1	R31001	Equine chorionic g
77	90	52.6	169	1	W33775	Equine chorionic g
78	90	52.6	169	1	W33777	Equine chorionic g
79	90	52.6	169	1	W33773	Equine chorionic g
80	90	52.6	169	1	W65110	Equine chorionic g
81	90	52.6	222	1	R86278	Single chain gonad
82	90	52.6	226	1	R86277	Single chain gonad
83	90	52.6	229	1	R86272	Single chain gonad
84	90	52.6	242	1	P60821	Sequence encoded b
85	90	52.6	265	1	W65111	Equine chorionic g
86	89	52.0	145	1	R15101	hCG/hLH chimera, D
87	89	52.0	145	1	W27687	Chorionic gonadotr
88	87	50.9	15	1	R34888	Human TSH residues
89	87	50.9	115	1	R15079	hCG/hFSH chimera,
90	87	50.9	118	1	R15077	hCG/hFSH chimera,
91	87	50.9	145	1	R15065	hCG/hFSH chimera,
92	87	50.9	145	1	R15074	hCG/hFSH chimera,
93	87	50.9	145	1	R15075	hCG/hFSH chimera,
94	87	50.9	145	1	R15076	hCG/hFSH chimera,
95	86	50.3	111	1	R15161	hCG deletion mutan
96	86	50.3	114	1	R31004	Modified hCG beta-
97	86	50.3	114	1	R31003	Modified hCG beta-
98	86	50.3	114	1	R31005	Modified hCG beta-
99	86	50.3	114	1	R88920	HCG analogue-G bet
100	86	50.3	114	1	R88921	HCG analogue-DG' b
101	86	50.3	114	1	R88922	HCG analogue-Q bet
102	86	50.3	115	1	R15078	hCG/hFSH chimera,
103	86	50.3	115	1	R15081	hCG/hFSH chimera,
104	86	50.3	115	1	R15083	hCG/hFSH chimera,
105	86	50.3	115	1	R15087	hCG/hFSH chimera,
106	86	50.3	116	1	R15162	hCG deletion mutan

107	86	50.3	116	1	R15172	hCG methionine sub	180	86	50.3	165	1	W99534	hCG-beta analogue
108	86	50.3	117	1	R31008	Modified hCG beta-	181	86	50.3	165	1	W99535	hCG/hFSH chimeric
109	86	50.3	117	1	R31007	Modified hCG beta-	182	86	50.3	165	1	W99536	hCG/hFSH chimeric
110	86	50.3	117	1	R88924	hCG analogue-GT be	183	86	50.3	165	1	W99538	hCG-beta analogue
111	86	50.3	117	1	R88925	hCG analogue-DGT b	184	86	50.3	165	1	W99539	hCG/hFSH chimeric
112	86	50.3	118	1	R15070	hCG/hFSH chimera,	185	86	50.3	165	1	W99541	hCG-beta analogue
113	86	50.3	118	1	R15071	hCG/hFSH chimera,	186	86	50.3	165	1	W99548	hCG-beta analogue
114	86	50.3	122	1	R15105	hCG/BLH chimera, D	187	86	50.3	165	1	W99528	Glycoprotein hormo
115	86	50.3	122	1	R15108	hCG/BLH chimera, D	188	86	50.3	165	1	W99512	Glycoprotein hormo
116	86	50.3	122	1	R15107	hCG/BLH chimera, D	189	86	50.3	165	1	W99512	Glycoprotein hormo
117	86	50.3	122	1	R15119	hCG/BLH chimera, A	190	86	50.3	165	1	W99513	Glycoprotein hormo
118	86	50.3	123	1	R15096	hCG/hFSH chimera,	191	86	50.3	165	1	W99515	Glycoprotein hormo
119	86	50.3	123	1	R15163	hCG deletion mutan	192	86	50.3	165	1	W99506	Glycoprotein hormo
120	86	50.3	124	1	R15095	hCG/hFSH chimera,	193	86	50.3	165	1	W99508	Glycoprotein hormo
121	86	50.3	128	1	R15166	hCG deletion mutan	194	86	50.3	165	1	W99509	Glycoprotein hormo
122	86	50.3	133	1	R15164	hCG deletion mutan	195	86	50.3	165	1	W99510	Glycoprotein hormo
123	86	50.3	137	1	R15168	hCG deletion mutan	196	86	50.3	165	1	W99511	Glycoprotein hormo
124	86	50.3	138	1	R15088	hCG/hFSH chimera,	197	86	50.3	181	1	W99511	Glycoprotein hormo
125	86	50.3	139	1	R15061	hCG/hFSH chimera,	198	86	50.3	234	1	R86266	Human CG beta-subu
126	86	50.3	139	1	R15167	hCG deletion mutan	199	86	50.3	234	1	R86265	Partially deglycos
127	86	50.3	141	1	R15168	hCG deletion mutan	200	86	50.3	234	1	R86264	Single chain gonad
128	86	50.3	142	1	W99526	Glycoprotein hormo	201	86	50.3	234	1	R86254	Single chain gonad
129	86	50.3	144	1	R15178	hCG histidine subs	202	86	50.3	234	1	R86248	Single chain gonad
130	86	50.3	144	1	R15093	hCG/hFSH chimera,	203	86	50.3	234	1	R86260	Partially deglycos
131	86	50.3	145	1	R15062	hCG/hFSH chimera,	204	86	50.3	237	1	R86263	Single chain gonad
132	86	50.3	145	1	R15063	hCG/hFSH chimera,	205	86	50.3	237	1	R86264	Partially deglycos
133	86	50.3	145	1	R15064	hCG/hFSH chimera,	206	86	50.3	237	1	R86251	Single chain gonad
134	86	50.3	145	1	R15068	hCG/hFSH chimera,	207	86	50.3	237	1	R86252	Single chain gonad
135	86	50.3	145	1	R15069	hCG/hFSH chimera,	208	86	50.3	265	1	R86247	Single chain gonad
136	86	50.3	145	1	R15072	hCG/hFSH chimera,	209	86	50.3	265	1	R86259	Partially deglycos
137	86	50.3	145	1	R15073	hCG/hFSH chimera,	210	86	50.3	307	1	W33358	TBP(20-161)/hCG-be
138	86	50.3	145	1	R15174	hCG histidine subs	211	86	50.3	335	1	W33360	TBP(20-190)/hCG-be
139	86	50.3	145	1	R15106	hCG/BLH chimera, D	212	85	49.7	139	1	P60601	Sequence of dog be
140	86	50.3	145	1	R15118	hCG/hFSH chimera, A	213	85	49.7	234	1	R86271	Single chain gonad
141	86	50.3	145	1	R15080	hCG/hFSH chimera,	214	84	49.1	145	1	R15067	hCG/hFSH chimera,
142	86	50.3	145	1	R15102	hCG/BLH chimera, D	215	83	48.5	15	1	R34887	Human TSH residues
143	86	50.3	145	1	R15089	hCG/hFSH chimera,	216	83	48.5	98	1	W50089	Human chorionic go
144	86	50.3	145	1	R15103	hCG/BLH chimera, D	217	83	48.5	98	1	W50054	Human chorionic go
145	86	50.3	145	1	R15099	hCG/BLH chimera, D	218	83	48.5	98	1	W47494	Human beta-hCG pro
146	86	50.3	145	1	R15098	hCG/BLH chimera, D	219	83	48.5	145	1	R15100	hCG/BLH chimera, D
147	86	50.3	145	1	R15097	hCG/BLH chimera, D	220	83	48.5	165	1	W33639	Human chorionic go
148	86	50.3	145	1	R15104	hCG/BLH chimera, D	221	83	48.5	165	1	W33637	Human chorionic go
149	86	50.3	145	1	R15109	hCG/ELH chimera, E	222	83	48.5	165	1	W47473	Human beta-hCG pro
150	86	50.3	145	1	R15110	hCG/ELH chimera, E	223	82	48.0	165	1	W99507	Glycoprotein hormo
151	86	50.3	145	1	R15111	hCG/ELH chimera, E	224	78	45.6	145	1	W27685	Chorionic gonadotr
152	86	50.3	145	1	R15112	hCG/ELH chimera, E	225	77	45.0	145	1	R15177	hCG histidine subs
153	86	50.3	145	1	R15113	hCG/ELH chimera, E	226	76	44.4	142	1	R54674	Fish gonadotropin
154	86	50.3	145	1	R15114	hCG/BLH chimera, A	227	76	44.4	145	1	R15176	hCG histidine subs
155	86	50.3	145	1	R15120	hCG/BLH chimera, A	228	76	44.4	234	1	R86275	Single chain gonad
156	86	50.3	145	1	R15121	hCG/BLH chimera, A	229	76	44.4	234	1	R86276	Single chain gonad
157	86	50.3	145	1	R15125	hCG/hFSH chimera,	230	76	44.4	234	1	R86270	Single chain gonad
158	86	50.3	145	1	R15084	hCG/hFSH chimera,	231	76	44.4	237	1	R86273	Single chain gonad
159	86	50.3	145	1	R15094	hCG/hFSH chimera,	232	76	44.4	237	1	R86274	Single chain gonad
160	86	50.3	145	1	R15169	hCG methionine sub	233	76	44.4	265	1	R86269	Single chain gonad
161	86	50.3	145	1	R15170	hCG methionine sub	234	70	40.9	88	1	W50090	Human chorionic go
162	86	50.3	145	1	R15171	hCG methionine sub	235	70	40.9	88	1	W50055	Human chorionic go
163	86	50.3	145	1	R15173	hCG histidine subs	236	70	40.9	88	1	W47495	Human beta-hCG pro
164	86	50.3	145	1	R30999	Human chorionic go	237	70	40.9	132	1	W69449	Human chorionic go
165	86	50.3	145	1	R31006	Modified hCG beta-	238	68.5	40.1	137	1	P94150	Fish gonadotropin
166	86	50.3	145	1	R88923	hCG analogue-D bet	239	68.5	40.1	137	1	R54673	Fish gonadotropin
167	86	50.3	145	1	W27688	Chorionic gonadotr	240	62	36.3	26	1	P67182	Bovine glial cell
168	86	50.3	145	1	W27681	Chorionic gonadotr	241	62	36.3	26	1	R86649	GGF 1 protease W8
169	86	50.3	145	1	W27682	Chorionic gonadotr	242	54	31.6	15	1	R34889	Human TSH residues
170	86	50.3	145	1	W27678	Chorionic gonadotr	243	49	28.7	15	1	P60822	Sequence encoded b
171	86	50.3	145	1	W27679	Chorionic gonadotr	244	48.5	28.4	256	1	W33357	TBP(20-161)/hCG-al
172	86	50.3	145	1	W27680	Chorionic gonadotr	245	47	27.5	453	1	W24250	Aquifex adenosyl-8
173	86	50.3	145	1	W31200	Human chorionic go	246	46.5	27.2	289	1	R23623	Tox2a gene product
174	86	50.3	145	1	W95520	Human chorionic go	247	46.5	27.2	289	1	W56745	Insect-specific pa
175	86	50.3	145	1	W99530	Human chorionic go	248	46	26.9	65	1	W30350	Fragment of inhibi
176	86	50.3	145	1	W93434	Human hCG beta-sub	249	46	26.9	134	1	P80018	Sequence of the 18
177	86	50.3	147	1	R15175	hCG insertion muta	250	46	26.9	351	1	P70202	Sequence of human
178	86	50.3	165	1	R15043	Human chorionic go	251	46	26.9	366	1	P60519	Sequence of human
179	86	50.3	165	1	W99533	Human chorionic go	252	46	26.9	366	1	P83167	Sequence of the 18

253	45.5	26.6	484	1	W19006	Feline herpes viru	326	40	23.4	328	1	R91224	Mouse pancreas G-p
254	45	26.3	493	1	W1068	Multiple sclerosis	327	40	23.4	477	1	P50585	Maize ribulose-bip
255	45	26.3	493	1	W9553	Protein encoded by	328	40	23.4	550	1	W47298	Human IDS. Treatme
256	44.5	26.0	83	1	R12181	Leukocyte elastase	329	40	23.4	550	1	W69786	Human iduronate-2-
257	44.5	26.0	285	1	W33359	TBP(20-190)/hCG-a1	330	40	23.4	553	1	W63703	Truncated rat rSK3
258	43	25.1	1128	1	R49994	Mouse carboxypepti	331	40	23.4	557	1	W63708	Truncated human hS
259	43	25.1	1128	1	W36817	Mouse E2A-binding	332	40	23.4	707	1	W47155	Extracellular doma
260	43	25.1	2254	1	R76949	ACCase. Plant acet	333	40	23.4	731	1	W96312	Human small conduc
261	42.5	24.9	351	1	W2967	Human serine carbo	334	40	23.4	732	1	W63715	Rat rSK3 protein.
262	42.5	24.9	398	1	W24252	Ammonifex degensii	335	40	23.4	736	1	W63717	Human hSK3 protein
263	42.5	24.9	476	1	W2966	Human serine carbo	336	40	23.4	750	1	R55097	Prostate-specific
264	42.5	24.9	476	1	Y13372	Amino acid sequenc	337	40	23.4	750	1	W02234	Prostate-specific
265	42.5	24.9	477	1	W2965	Human serine carbo	338	40	23.4	751	1	W31524	Prostate-specific
266	42.5	24.9	764	1	W81784	P. putida cis/trans	339	40	23.4	758	1	R86998	Enterohaemorrhagic
267	42	24.6	581	1	Y03831	L-galactonolactone	340	40	23.4	848	1	Y07733	Human thyroid pero
268	42	24.6	680	1	R07453	First open reading	341	40	23.4	875	1	W37098	Murine endothelial
269	42	24.6	680	1	W46516	Saccharomyces cere	342	40	23.4	1266	1	W48721	Murine cytoplasmic
270	42	24.6	767	1	W98606	H. pylori GHPO 147	343	40	23.4	1830	1	R44504	Urea amide lyase.
271	42	24.6	1686	1	W70991	Human class II p13	344	40	23.4	3066	1	W36178	Murine Ataxia-tela
272	42	24.6	1726	1	W38756	Phosphatidyl inosi	345	39.5	23.1	116	1	W56494	Tobacco leaf curl
273	41.5	24.3	64	1	Y11660	Human 5' EST secre	346	39.5	23.1	125	1	R63060	Rat P1A2 type II.
274	41.5	24.3	171	1	Y11333	S. pneumoniae prot	347	39.5	23.1	128	1	W26473	RFLV glycoprotein
275	41.5	24.3	191	1	R63331	Hepatitis C virus	348	39.5	23.1	145	1	R63053	HPLA2-8. Novel typ
276	41.5	24.3	193	1	R63371	Hepatitis C virus	349	39.5	23.1	193	1	R63376	Hepatitis C virus
277	41	24.0	87	1	R15192	hCG alpha subunit	350	39.5	23.1	193	1	R63370	Hepatitis C virus
278	41	24.0	92	1	P92146	Beta subunit of ur	351	39.5	23.1	325	1	W20992	H. pylori outer me
279	41	24.0	92	1	R15193	hCG alpha subunit	352	39.5	23.1	325	1	W20534	H. pylori cell env
280	41	24.0	92	1	R15194	hCG alpha subunit	353	39.5	23.1	325	1	W24683	H. pylori outer me
281	41	24.0	92	1	R15181	hCG/bCG alpha subu	354	39.5	23.1	436	1	W93151	C. saccharolyticum
282	41	24.0	92	1	W95519	Human chorionic go	355	39.5	23.1	486	1	W76822	Human Tcf-1B prote
283	41	24.0	92	1	W95529	Human chorionic go	356	39.5	23.1	511	1	W76823	Human Tcf-1B prote
284	41	24.0	93	1	R15196	hCG alpha subunit	357	39.5	23.1	513	1	R14304	Insect steroid rec
285	41	24.0	96	1	R15195	hCG alpha subunit	358	39.5	23.1	513	1	R48065	D. melanogaster ul
286	41	24.0	101	1	R86257	Deglycosylated gon	359	39.5	23.1	513	1	W18876	Insect XR2C recept
287	41	24.0	116	1	R10095	Engineered human a	360	39.5	23.1	513	1	W31320	D. melanogaster XR
288	41	24.0	116	1	R99417	Dimeric glycoprote	361	39.5	23.1	617	1	R43393	JM haemagglutinin.
289	41	24.0	116	1	W31665	Secreted protein h	362	39.5	23.1	767	1	Y11030	H. pylori ORF Olcp
290	41	24.0	116	1	W99537	hCG analogue alpha	363	39.5	23.1	2466	1	R71498	Human protein tyro
291	41	24.0	116	1	W99542	hCG analogue alpha	364	39.5	23.1	2466	1	W75999	Intracellular prot
292	41	24.0	116	1	W99544	GPH analogue alpha	365	39	22.8	74	1	W74817	Human secreted pro
293	41	24.0	116	1	W99531	Human chorionic go	366	39	22.8	96	1	R15185	hCG/bCG alpha subu
294	41	24.0	116	1	W99491	Glycoprotein hormo	367	39	22.8	96	1	R15189	hCG/bCG alpha subu
295	41	24.0	116	1	W99492	Glycoprotein hormo	368	39	22.8	96	1	R15190	hCG/bCG alpha subu
296	41	24.0	116	1	W99493	Glycoprotein hormo	369	39	22.8	111	1	R82065	Hepatitis GB virus
297	41	24.0	116	1	W99494	Glycoprotein hormo	370	39	22.8	112	1	W77578	Staphylococcus aur
298	41	24.0	116	1	W99495	Glycoprotein hormo	371	39	22.8	146	1	W08368	Mouse P1A2s (wildt
299	41	24.0	116	1	W99484	Glycoprotein hormo	372	39	22.8	280	1	W97635	Xenopus Frazzled p
300	41	24.0	116	1	W99485	Glycoprotein hormo	373	39	22.8	281	1	W97634	Xenopus Frazzled p
301	41	24.0	116	1	W99486	Glycoprotein hormo	374	39	22.8	311	1	Y05316	S. aureus protein
302	41	24.0	116	1	W99487	Glycoprotein hormo	375	39	22.8	334	1	R52578	Glucanase of Horde
303	41	24.0	116	1	W99488	Glycoprotein hormo	376	39	22.8	338	1	R27667	Soya beta-1,3-gluc
304	41	24.0	116	1	W99489	Glycoprotein hormo	377	39	22.8	338	1	R56862	Soya beta-1,3-gluc
305	41	24.0	116	1	W99490	Glycoprotein hormo	378	39	22.8	370	1	R27666	Soya beta-1,3-gluc
306	41	24.0	116	1	W99483	Glycoprotein hormo	379	39	22.8	370	1	R27668	35S CaMV promoter-
307	41	24.0	191	1	R92942	Hepatitis C virus	380	39	22.8	421	1	R97599	Adenovirus 8 subty
308	41	24.0	306	1	R67974	Glucanase GII. New	381	39	22.8	438	1	W18209	Staphylococcus aur
309	41	24.0	656	1	W04627	Mouse receptor int	382	39	22.8	473	1	R21977	RuBPCase large sub
310	41	24.0	656	1	W80994	Human receptor int	383	39	22.8	624	1	W48789	Thyroid peroxidase
311	41	24.0	1964	1	W95557	Mus musculus notch	384	39	22.8	626	1	W48784	Thyroid peroxidase
312	40.5	23.7	266	1	R22665	Protein used to ra	385	39	22.8	646	1	W15470	Hantaan virus poly
313	40.5	23.7	554	1	W64538	Human liver cell c	386	39	22.8	689	1	W48787	Thyroid peroxidase
314	40.5	23.7	925	1	R79148	Human insulin rece	387	39	22.8	740	1	W48790	Thyroid peroxidase
315	40	23.4	17	1	R06091	Immunoreactive pep	388	39	22.8	768	1	W36815	Human E2A-binding
316	40	23.4	68	1	R15197	hCG/bCG alpha subu	389	39	22.8	768	1	W36823	Human E2A-binding
317	40	23.4	89	1	W38750	Streptococcus pneu	390	39	22.8	784	1	W48783	Thyroid peroxidase
318	40	23.4	96	1	R15182	hCG/bCG alpha subu	391	39	22.8	845	1	W36816	Human E2A-binding
319	40	23.4	96	1	R15183	hCG/bCG alpha subu	392	39	22.8	852	1	W48782	Thyroid peroxidase
320	40	23.4	96	1	R15188	hCG/bCG alpha subu	393	39	22.8	881	1	W48791	Thyroid peroxidase
321	40	23.4	96	1	R15191	hCG/bCG alpha subu	394	39	22.8	924	1	R35445	Human TPO lacking
322	40	23.4	116	1	W99532	hCG analogue alpha	395	39	22.8	930	1	W75426	Human aggrecan deg
323	40	23.4	116	1	W99540	hCG analogue alpha	396	39	22.8	933	1	R32875	Human TPO. Peptide
324	40	23.4	133	1	W77684	4-Methyl-5(beta-hy	397	39	22.8	933	1	R44615	Human thyroic pero
325	40	23.4	218	1	R79164	Partial sequence o	398	39	22.8	933	1	R75669	Human thyroid pero

399	39	22.8	948	1	W48781	Thyroid peroxidase	472	37	21.6	216	1	W59897	Amino acid sequenc
400	39	22.8	1135	1	R50037	Hantaan virus G1/	473	37	21.6	216	1	W59903	Amino acid sequenc
401	39	22.8	1233	1	W34536	Nudaurelia beta vi	474	37	21.6	216	1	W59899	Amino acid sequenc
402	39	22.8	1233	1	W41935	Nudaurelia beta-1i	475	37	21.6	219	1	W59888	Amino acid sequenc
403	39	22.8	1370	1	R91430	Rat type I insulin	476	37	21.6	219	1	W59904	Amino acid sequenc
404	39	22.8	1422	1	R82068	Hepatitis GB virus	477	37	21.6	219	1	W59902	Amino acid sequenc
405	39	22.8	2647	1	W19349	Human filamin. Hum	478	37	21.6	220	1	W59905	Amino acid sequenc
406	39	22.8	2864	1	R82072	Hepatitis GB virus	479	37	21.6	223	1	W59896	Amino acid sequenc
407	38.5	22.5	94	1	Y12395	Human 5' EST secre	480	37	21.6	223	1	W59898	Amino acid sequenc
408	38.5	22.5	178	1	W11994	Herpes virus DNA P	481	37	21.6	251	1	R15611	SP-C from PC210SP-
409	38.5	22.5	192	1	R69674	Hepatitis C virus	482	37	21.6	286	1	R1820	Human immunodefici
410	38.5	22.5	192	1	R89545	Hepatitis C virus	483	37	21.6	288	1	W14450	CarF gene product.
411	38.5	22.5	193	1	R63369	Hepatitis C virus	484	37	21.6	289	1	W77303	Amino acid sequenc
412	38.5	22.5	229	1	W49054	Hepatitis C virus	485	37	21.6	305	1	Y03181	Topoisomerase II b
413	38.5	22.5	283	1	R10975	Human Borna disease	486	37	21.6	315	1	W61616	Clone HMAR70 of I
414	38.5	22.5	443	1	W32110	Polyhydroxyalkanoa	487	37	21.6	331	1	W22011	Glucan elicitor re
415	38.5	22.5	537	1	Y03839	Human extracellular	488	37	21.6	331	1	W22011	Glucan elicitor re
416	38.5	22.5	837	1	W11995	Chimeric transcrip	489	37	21.6	358	1	R98919	Sequence in a lowh
417	38.5	22.5	884	1	W59778	Herpes virus DNA p	490	37	21.6	359	1	R28274	Glucanase protein
418	38.5	22.5	884	1	W71375	Amino acid sequenc	491	37	21.6	366	1	W97731	Sequence in a lowh
419	38.5	22.5	896	1	R63533	Yeast telomerase c	492	37	21.6	374	1	W69405	Soybean diamnopl m
420	38.5	22.5	1608	1	R98619	Human HT-1376 cell	493	37	21.6	392	1	W59993	Lactic acid oxidas
421	38.5	22.5	1711	1	R98605	Borna disease viru	494	37	21.6	403	1	Y03182	Sialoadhesin famil
422	38.5	22.5	2386	1	W13153	Borna disease viru	495	37	21.6	448	1	R11350	Topoisomerase II b
423	38.5	22.5	4544	1	R47861	S. pombe Rad3 poly	496	37	21.6	458	1	W64391	Cytochrome enzyme
424	38.5	22.5	4544	1	R60517	Alpha 2-Macroglobu	497	37	21.6	467	1	W49908	A. thaliana SGT pr
425	38	22.2	267	1	R03346	Human alpha-2-MR.	498	37	21.6	473	1	W59992	Sialoadhesin famil
426	38	22.2	289	1	R81054	VP2 sequence for H	499	37	21.6	473	1	R66610	Human cdc25C gene
427	38	22.2	293	1	R92221	Sequence of rhinov	500	37	21.6	473	1	R98214	Human CDC25C phosph
428	38	22.2	333	1	W88345	Glycine-N-methyl t	501	37	21.6	538	1	W59887	Amino acid sequenc
429	38	22.2	383	1	P70291	Salmonella enteric	502	37	21.6	559	1	P94403	Human endogenous r
430	38	22.2	426	1	P40254	Sequence encoded b	503	37	21.6	561	1	P70020	Analogue of amino
431	38	22.2	456	1	W69437	Human IgD. Novel i	504	37	21.6	561	1	R91065	Sequence of tissue
432	38	22.2	470	1	R34473	Rat phospholipase	505	37	21.6	562	1	R05388	Rat interleukin-1
433	38	22.2	471	1	R79163	Encoded by Hepatit	506	37	21.6	562	1	P70285	Thrombolytic prote
434	38	22.2	572	1	W68487	Partial sequence o	507	37	21.6	562	1	R91064	Sequence encoded b
435	38	22.2	699	1	W81840	Mouse ULIP-3 prote	508	37	21.6	567	1	R88427	Human interleukin-
436	38	22.2	769	1	P80836	Human LFA-1 beta c	509	37	21.6	582	1	W01103	Arabidopsis thalia
437	38	22.2	769	1	R07113	Beta subunit Of hu	510	37	21.6	582	1	W01104	N. benthamiana phy
438	38	22.2	769	1	R24256	Recombinant beta-s	511	37	21.6	582	1	W01105	N. benthamiana phy
439	38	22.2	769	1	R80108	Beta-subunit CD18	512	37	21.6	637	1	W64388	N. tabacum phytoen
440	38	22.2	786	1	W86349	LFA-1 beta subunit	513	37	21.6	674	1	W64389	A. thaliana SGT pr
441	38	22.2	819	1	R10648	Human DNAX toll-1i	514	37	21.6	811	1	W93952	A. thaliana protei
442	38	22.2	821	1	R99579	Chicken basic fibr	515	37	21.6	816	1	R71111	Human regulatory m
443	38	22.2	1051	1	W67641	Calpain large subu	516	37	21.6	824	1	R04575	Spinocerebellar at
444	38	22.2	1827	1	W74090	A serine/threonine	517	37	21.6	825	1	R11254	Derived amino acid
445	38	22.2	2164	1	R80131	Human hsi protein	518	37	21.6	825	1	W13499	Human IL-4 recepto
446	38	22.2	2164	1	R81045	Peptides translate	519	37	21.6	825	1	W48314	Human interleukin-
447	37.5	21.9	193	1	R63363	Sequence of the vi	520	37	21.6	825	1	W60668	Human interleukin-
448	37.5	21.9	193	1	R63362	Hepatitis C virus	521	37	21.6	825	1	W73470	Human interleukin-
449	37.5	21.9	193	1	R33699	Hepatitis C virus	522	37	21.6	825	1	W93004	Human interleukin-
450	37.5	21.9	553	1	W68489	Sequence of soybea	523	37	21.6	840	1	W63112	Human IL-4 recepto
451	37.5	21.9	572	1	W68488	Human partial ULIP	524	37	21.6	847	1	W63108	Protein encoded by
452	37.5	21.9	613	1	W22051	Mouse ULIP-4 prote	525	37	21.6	918	1	P60723	Protein encoded by
453	37.5	21.9	613	1	W97097	DNA polymerase 220	526	37	21.6	1050	1	W67642	Sequence of extrac
454	37.5	21.9	657	1	W71851	Pfu DNA polymerase	527	37	21.6	1073	1	R28821	A serine/threonine
455	37.5	21.9	1211	1	W24560	Mouse membrane typ	528	37	21.6	1073	1	R28821	Alpha 6A integrin
456	37.5	21.9	1404	1	R26049	Presenilin-interac	529	37	21.6	1091	1	R28822	Alpha subunit of i
457	37.5	21.9	2322	1	W15566	MSF precursor. New	530	37	21.6	1240	1	W03665	Alpha 6B integrin
458	37	21.6	17	1	R06094	Melanoma-associate	531	37	21.6	1385	1	R20066	I2C-2 protein conf
459	37	21.6	104	1	W74983	Immunoreactive pep	532	37	21.6	1385	1	R28889	B. thuringiensis to
460	37	21.6	126	1	R90543	Human secreted pro	533	37	21.6	1385	1	R29026	Toxin 17A. New Bac
461	37	21.6	132	1	W74984	pUG4-5-CDK-Bp clon	534	37	21.6	1385	1	R28803	Bacillus thuringie
462	37	21.6	143	1	R28122	Human secreted pro	535	37	21.6	1385	1	R29516	Bacillus thuringie
463	37	21.6	154	1	W74837	Alpha-amylase/tyrp	536	37	21.6	1385	1	R58631	BT toxin 17A. Nema
464	37	21.6	160	1	R78652	Human secreted pro	537	37	21.6	1385	1	R44201	Bacillus thuringie
465	37	21.6	161	1	Y11868	Human skeletal mus	538	37	21.6	1385	1	R76112	PS17a acaride-acti
466	37	21.6	162	1	R78693	Human 5' EST secre	539	37	21.6	1385	1	W13884	17A toxin. Bacillu
467	37	21.6	188	1	R77033	Rat skeletal muscl	540	37	21.6	4302	1	W00870	Polycystic kidney
468	37	21.6	188	1	W88277	Mammalian mast cel	541	37	21.6	4302	1	W23830	Human PKD1 protein
469	37	21.6	189	1	W59901	Rat mast cell func	542	37	21.6	4302	1	W33396	Human PKD1 polypep
470	37	21.6	192	1	W59889	Amino acid sequenc	543	37	21.6	4303	1	R90302	Polycystic kidney
471	37	21.6	192	1	W59900	Amino acid sequenc	544	37	21.6	4339	1	R75916	Polycystic kidney

545	4339	1	R87539	Polycystic kidney	618	240	1	W32255	Human apoptosis in
546	37	21.6	1	R05709	Heat-resistant RNA	619	240	1	Human Fas ligand-1
547	36.5	21.3	1	R26474	KSHV glycoprotein	620	240	1	Tumour necrosis fa
548	36.5	21.3	1	R96540	Hepatitis C virus	621	240	1	L. ivanovii strain
549	36.5	21.3	1	R96541	Hepatitis C virus	622	284	1	L. seeligeri strain
550	36.5	21.3	1	R96542	Hepatitis C virus	623	284	1	L. welshimeri strai
551	36.5	21.3	1	R63368	Hepatitis C virus	624	284	1	L. monocytogenes d
552	36.5	21.3	1	R63361	Hepatitis C virus	625	284	1	L. monocytogenes s
553	36.5	21.3	1	W22440	Human neurogenic d	626	284	1	L. innocua strain
554	36.5	21.3	1	W71016	Human neuroD3 prot	627	323	1	E. coli O157 antiG
555	36.5	21.3	1	R94617	3-hydroxybutyric a	628	327	1	H. pylori ORF 02ae
556	36.5	21.3	1	R95968	Nucleic acid recog	629	327	1	H. pylori GHPO 146
557	36.5	21.3	1	W29816	Mammalian AMPK bet	630	328	1	Human placenta pur
558	36.5	21.3	1	R90544	pG4-5-CDK-BP clon	631	337	1	Human Pax9 gene pr
559	36.5	21.3	1	W15760	Cotton fibrous tis	632	343	1	H. pylori ORF hplp
560	36.5	21.3	1	R12346	Toxoplasma gondii	633	352	1	Murine liver activ
561	36.5	21.3	1	W01731	T. gondii antigen	634	366	1	DAZ protein. DAZ g
562	36.5	21.3	1	W19990	Type I, p80 IL-1-r	635	378	1	W41508
563	36.5	21.3	1	W64555	Human osteocarcin	636	389	1	Y01090
564	36.5	21.3	1	R48701	G-protein coupled	637	398	1	W03518
565	36.5	21.3	1	W02673	G-protein coupled	638	404	1	R83232
566	36.5	21.3	1	R52635	RaEPV p50 spindle	639	411	1	R90574
567	36.5	21.3	1	W02206	Insect poxvirus 38	640	411	1	R90576
568	36.5	21.3	1	W95490	M. tuberculosis RP	641	411	1	R86810
569	36.5	21.3	1	W19111	Human tumour susce	642	411	1	W16481
570	36.5	21.3	1	W31740	Delta-12 desaturase	643	416	1	W54098
571	36.5	21.3	1	W93424	Human NHR5 protein	644	422	1	R99090
572	36.5	21.3	1	W04217	Rat dorsal root ga	645	431	1	R90575
573	36.5	21.3	1	W76437	Human p53 regulate	646	431	1	R97293
574	36.5	21.3	1	R12353	Toxoplasma gondii	647	435	1	R13907
575	36.5	21.3	1	W01732	T. gondii antigen	648	446	1	R11837
576	36.5	21.3	1	W01732	S. pombe squalene	649	446	1	R72484
577	36.5	21.3	1	R82080	Active domain of h	650	448	1	R39467
578	36.5	21.3	1	R39635	Human 3-hydroxy-3-	651	453	1	W49906
579	36.5	21.3	1	W26524	Epsilon-rhodomycin	652	457	1	R82931
580	36.5	21.3	1	R22925	Truncated hamster	653	459	1	R21312
581	36.5	21.3	1	R91532	Phosphatidylinosit	654	459	1	R26175
582	36.5	21.3	1	W77772	Huma glutamine:fru	655	462	1	R07624
583	36.5	21.3	1	W26475	KSHV glycoprotein	656	462	1	W11237
584	36.5	21.3	1	W22222	HSV-2 strain SB5 C	657	476	1	W02616
585	36.5	21.3	1	W72023	HSV-2 strain SB5 C	658	509	1	R25202
586	36.5	21.3	1	W72126	HSV-2 strain SB5 C	659	509	1	R47348
587	36.5	21.3	1	W72223	HSV-2 strain SB5 C	660	509	1	W44280
588	36.5	21.3	1	R58609	Hamster HMG-CoA re	661	510	1	R39469
589	36.5	21.3	1	W22863	Bacillus stearothe	662	511	1	W12369
590	36.5	21.3	1	W44320	Bacillus thuringie	663	516	1	W15286
591	36.5	21.3	1	W93601	Tomato Xa21 clone	664	532	1	R86957
592	36	21.1	1	R06088	Immunoreactive pep	665	533	1	R39488
593	36	21.1	1	P71015	Sequence of fragme	666	533	1	R72483
594	36	21.1	20	W47505	Human beta-hCG pro	667	551	1	P91400
595	36	21.1	20	W55888	Rat PC12. Telomera	668	553	1	R86956
596	36	21.1	58	R69522	Mouse corticotropi	669	560	1	R15051
597	36	21.1	70	R97295	Mouse CRF receptor	670	561	1	P61363
598	36	21.1	92	R15179	hCG/bCG alpha subu	671	582	1	W67639
599	36	21.1	111	W75889	Peptide inhibitor	672	663	1	R24101
600	36	21.1	136	W55889	Rat plasmid RaPC53	673	684	1	W49873
601	36	21.1	138	R98551	H. pylori GHPO 161	674	772	1	R23873
602	36	21.1	138	W98442	H. pylori GHPO 519	675	772	1	W09643
603	36	21.1	139	R86958	Human ribonuclease	676	823	1	W04358
604	36	21.1	163	W88487	Hepatitis C virus	677	836	1	R86954
605	36	21.1	192	R69681	Hepatitis C virus	678	844	1	R86952
606	36	21.1	192	R69653	Hepatitis C virus	679	844	1	R86952
607	36	21.1	192	R69644	Hepatitis C virus	680	897	1	Y07092
608	36	21.1	192	R89515	Hepatitis C virus	681	934	1	R14402
609	36	21.1	192	R89524	Hepatitis C virus	682	934	1	R15048
610	36	21.1	192	R89552	Hepatitis C virus	683	935	1	R15050
611	36	21.1	193	R68612	Saccharomyces cere	684	948	1	R14403
612	36	21.1	201	R98618	Borna disease viru	685	948	1	R15049
613	36	21.1	201	W49043	Human Borna disease	686	1037	1	W67643
614	36	21.1	201	W49044	Human Borna disease	687	1146	1	W96322
615	36	21.1	201	W49045	Human Borna disease	688	1153	1	R55764
616	36	21.1	229	P93524	Complete sequence	689	1153	1	R63206
617	36	21.1	236	1	Rat corticotropin-	690	1153	1	R88464

691	36	21.1	1153	1	W36113	Human inducible ni	764	35	20.5	132	1	W07861	(DSM 10103) human
692	36	21.1	1188	1	R94622	HIV-1 gp160-Nr5 p5	765	35	20.5	133	1	P70865	Sequence of S5 sub
693	36	21.1	1337	1	R63123	IGF-1 receptor 943	766	35	20.5	135	1	W47284	Papillomavirus maj
694	36	21.1	1337	1	R63124	IGF-1 receptor 950	767	35	20.5	138	1	R63046	HP1A-10. Novel ty
695	36	21.1	1337	1	R63125	IGF-1 receptor 957	768	35	20.5	142	1	W82676	Soil bacteria type
696	36	21.1	1337	1	R63126	IGF-1 receptor 100	769	35	20.5	145	1	P81003	Sequence encoding
697	36	21.1	1367	1	R60795	Human IFG-1 recept	770	35	20.5	153	1	R20099	Wound-inducible po
698	36	21.1	1367	1	R63122	IGF-1 receptor. Re	771	35	20.5	153	1	Y07965	Human secreted pro
699	36	21.1	1367	1	R91429	Human type I insul	772	35	20.5	154	1	W76227	Bacterial periplas
700	36	21.1	1367	1	R95244	IGF-1 receptor. In	773	35	20.5	191	1	R92954	Hepatitis C virus
701	36	21.1	1367	1	W37692	Human insulin-like	774	35	20.5	192	1	R69642	Hepatitis C virus
702	36	21.1	1367	1	W54876	Homo sapiens IGF-1	775	35	20.5	192	1	R69643	Hepatitis C virus
703	36	21.1	2050	1	W73499	Von Willebrand fac	776	35	20.5	192	1	R69655	Hepatitis C virus
704	36	21.1	2050	1	W90117	Human mature von W	777	35	20.5	192	1	R69656	Hepatitis C virus
705	36	21.1	2629	1	W55885	Rat telomerase. Te	778	35	20.5	192	1	R69648	Hepatitis C virus
706	36	21.1	2629	1	W61348	Mouse telomerase R	779	35	20.5	192	1	R89527	Hepatitis C virus
707	36	21.1	3023	1	P60462	Sequence of human	780	35	20.5	192	1	R89513	Hepatitis C virus
708	36	21.1	3023	1	R94462	Hepatitis C virus	781	35	20.5	192	1	R89514	Hepatitis C virus
709	35.5	20.8	32	1	W74888	Human secreted pro	782	35	20.5	192	1	R89519	Hepatitis C virus
710	35.5	20.8	60	1	Y02727	Human secreted pro	783	35	20.5	192	1	R89526	Hepatitis C virus
711	35.5	20.8	128	1	W26476	RFHV glycoprotein	784	35	20.5	195	1	W20378	Helicobacter pylor
712	35.5	20.8	170	1	R94977	Rex region gene pr	785	35	20.5	195	1	W24047	H. pylori outer me
713	35.5	20.8	172	1	W86324	Kidney injury asso	786	35	20.5	220	1	W23681	French bean polyph
714	35.5	20.8	188	1	R79915	Human trypsin inh	787	35	20.5	223	1	W72136	HSV-2 strain S85 C
715	35.5	20.8	191	1	R63332	Hepatitis C virus	788	35	20.5	235	1	W48312	Mouse BCL-x gamma
716	35.5	20.8	191	1	R92962	Hepatitis C virus	789	35	20.5	238	1	W72066	HSV-2 strain S85 C
717	35.5	20.8	191	1	R92965	Hepatitis C virus	790	35	20.5	244	1	R44532	Zyme APP-cleaving
718	35.5	20.8	193	1	R63365	Hepatitis C virus	791	35	20.5	244	1	W22985	Human serine prote
719	35.5	20.8	198	1	R79914	Trypsin inhibitory	792	35	20.5	244	1	W51006	Protease M, a nove
720	35.5	20.8	241	1	P93282	Protein encoded by	793	35	20.5	248	1	W38544	Streptococcus pneu
721	35.5	20.8	292	1	R76952	Wild-type E.coli d	794	35	20.5	248	1	W83370	Streptococcus pneu
722	35.5	20.8	292	1	W18809	Dihydropicolinic a	795	35	20.5	250	1	W61619	Streptococcus pneu
723	35.5	20.8	292	1	W18808	Dihydropicolinic a	796	35	20.5	256	1	W25970	ORF 2 protein of P
724	35.5	20.8	292	1	W18807	Dihydropicolinic a	797	35	20.5	256	1	W25954	ORF 2 protein of P
725	35.5	20.8	292	1	W01783	Wild type dihydrop	798	35	20.5	274	1	W49837	Amino acid sequenc
726	35.5	20.8	417	1	R52606	Human squalene syn	799	35	20.5	277	1	R41432	PT-NANBH virus str
727	35.5	20.8	417	1	W01739	Human squalene syn	800	35	20.5	278	1	R12596	Antigenic portion
728	35.5	20.8	474	1	W59049	Variant of GABA-A	801	35	20.5	300	1	W20713	H. pylori outer me
729	35.5	20.8	506	1	W59048	GABA-A receptor ep	802	35	20.5	318	1	R96563	Hepatitis C virus
730	35.5	20.8	506	1	W61045	Human GABA recepto	803	35	20.5	321	1	W83375	Streptococcus pneu
731	35.5	20.8	685	1	R89115	Bombyx mori (pro)p	804	35	20.5	343	1	P70986	Streptococcus pneu
732	35.5	20.8	878	1	Y00868	S. tuberosum isoam	805	35	20.5	349	1	P70463	Sequence of gpi en
733	35.5	20.8	942	1	Y00871	Original S. tubero	806	35	20.5	352	1	R89729	Transferring growt
734	35.5	20.8	959	1	R76066	Yeast MSH1 protein	807	35	20.5	364	1	W22999	Canine herpesvirus
735	35.5	20.8	971	1	R76070	Yeast MSH1 protein	808	35	20.5	364	1	W72659	Canine herpes viru
736	35.5	20.8	1288	1	W26328	Mouse alpha-1 coll	809	35	20.5	368	1	R69570	H. pylori GHPO 83
737	35.5	20.8	2913	1	W00384	Plasmodium falcipa	810	35	20.5	384	1	W98586	Alfalfa beta-1,3-g
738	35	20.5	26	1	Y12809	SRC SH3 domain-bin	811	35	20.5	394	1	R13750	H. pylori GHPO 83
739	35	20.5	26	1	Y12809	Human 5' EST secre	812	35	20.5	394	1	R45935	GDP-Fuc:[beta-D-Ga
740	35	20.5	38	1	R65153	Specific TSAR 12'	813	35	20.5	394	1	W13639	A glycosyltransfer
741	35	20.5	42	1	Y11437	Human 5' EST secre	814	35	20.5	400	1	W17831	Murine alpha(1,3)-
742	35	20.5	42	1	Y11632	Human 5' EST secre	815	35	20.5	432	1	R92813	Human mevalonate p
743	35	20.5	52	1	R51391	Branched peptide H	816	35	20.5	441	1	R99091	Murine interleukin
744	35	20.5	53	1	R422996	Canine herpesvirus	817	35	20.5	442	1	W83369	Murine Etl-2 gene
745	35	20.5	53	1	W72655	Canine herpes viru	818	35	20.5	446	1	W85122	Streptococcus pneu
746	35	20.5	54	1	R51392	Branched peptide H	819	35	20.5	446	1	W83374	A delta-6 desatura
747	35	20.5	54	1	R51393	Branched peptide H	820	35	20.5	448	1	R98455	Streptococcus pneu
748	35	20.5	59	1	R51394	Branched peptide H	821	35	20.5	448	1	W67471	Borage delta-6-des
749	35	20.5	59	1	R51395	Branched peptide H	822	35	20.5	448	1	W98130	Borage delta-6 des
750	35	20.5	60	1	R52243	Subtilase BSBP fr	823	35	20.5	455	1	W97860	Cat flea glutamate
751	35	20.5	87	1	W45447	TGF-beta-like clon	824	35	20.5	456	1	W05246	Drosophila glutama
752	35	20.5	89	1	W37411	Tsetse thrombin in	825	35	20.5	457	1	W98284	H. pylori GHPO 126
753	35	20.5	89	1	W98573	H. pylori GHPO 25	826	35	20.5	476	1	W44827	H. insolens cellul
754	35	20.5	92	1	R51580	hCG/bCG alpha subu	827	35	20.5	476	1	W44823	Humicola insolens
755	35	20.5	92	1	Y07926	Human secreted pro	828	35	20.5	483	1	W19351	Filamin-like beta
756	35	20.5	96	1	R51587	hCG/bCG alpha subu	829	35	20.5	507	1	W00290	Mature Penicillin
757	35	20.5	100	1	R80237	Fragment of human	830	35	20.5	532	1	W00291	Full length Penici
758	35	20.5	110	1	W73327	Bipolar affective	831	35	20.5	533	1	R81549	High affinity pota
759	35	20.5	116	1	P81906	N-terminal of inh	832	35	20.5	545	1	W30708	Murine ubiquitin-s
760	35	20.5	118	1	W99582	Solid phase sequen	833	35	20.5	545	1	Y06860	Mutant luciferase
761	35	20.5	120	1	P96016	Pertussis toxin su	834	35	20.5	546	1	Y06858	Mutant luciferase
762	35	20.5	120	1	R94970	Bordetella pertuss	835	35	20.5	546	1	Y06859	Mutant luciferase
763	35	20.5	124	1	W79945	Staphylococcus aur	836	35	20.5	546	1	Y06857	Mutant luciferase

837	35	20.5	556	1	W27598	Human fibulin type	910	35	20.5	922	1	R55531	Human NMDA R1d rec
838	35	20.5	566	1	R11148	Fibulin A. Purifie	911	35	20.5	922	1	R66046	Human NMDA recepto
839	35	20.5	566	1	R13228	Endoglucanase enco	912	35	20.5	922	1	R66043	Human N-methyl-D-a
840	35	20.5	567	1	R76235	Bovine foetal hear	913	35	20.5	922	1	R85580	Human N-methyl-D-a
841	35	20.5	572	1	R131855	Mycobacterium tube	914	35	20.5	922	1	W85577	Human N-methyl-D-a
842	35	20.5	579	1	W63707	Human hsk2 protein.	915	35	20.5	925	1	R21606	G6 amylase. G6 amy
843	35	20.5	580	1	W63702	Rat rSK2 protein.	916	35	20.5	929	1	R66056	Human NMDA recepto
844	35	20.5	582	1	W01106	N. tabacum phytoen	917	35	20.5	929	1	W87301	N-methyl-D-asparta
845	35	20.5	585	1	P70143	BAR1 gene product.	918	35	20.5	930	1	R41431	PT-NANBH virus BHC
846	35	20.5	587	1	R13383	Barrier protein. D	919	35	20.5	938	1	R55532	Human NMDA R1a rec
847	35	20.5	587	1	R20109	BAR1 barrier prote	920	35	20.5	938	1	R66035	Human N-methyl-D-a
848	35	20.5	590	1	W37130	Hepatitits C virus	921	35	20.5	938	1	W85572	Human N-methyl-D-a
849	35	20.5	601	1	R11149	Fibulin B. Purifie	922	35	20.5	943	1	R66054	Human NMDA recepto
850	35	20.5	601	1	W27599	Human fibulin type	923	35	20.5	943	1	W85588	Human N-methyl-D-a
851	35	20.5	603	1	R12600	PT-NANBH viral str	924	35	20.5	947	1	W42402	Amino acid sequenc
852	35	20.5	610	1	R94649	Human Fas-associat	925	35	20.5	947	1	W81564	Human NF-kB-induci
853	35	20.5	610	1	W65770	Soybean phragmopla	926	35	20.5	947	1	W81561	Human NF-kB-induci
854	35	20.5	617	1	R21578	HCV CKS-ENV - pHCV	927	35	20.5	947	1	W82497	Human NIK protein.
855	35	20.5	617	1	R33587	HCV CKS-ENV recomb	928	35	20.5	959	1	R66048	Human NMDA recepto
856	35	20.5	617	1	R33607	HCV CKS-ENV recomb	929	35	20.5	959	1	W85582	Human N-methyl-D-a
857	35	20.5	617	1	R33645	HCV CKS-ENV recomb	930	35	20.5	976	1	R66047	Human NMDA recepto
858	35	20.5	661	1	W68009	Human soluble vasc	931	35	20.5	976	1	W85581	Human N-methyl-D-a
859	35	20.5	678	1	W72244	HSV-2 strain SB5 C	932	35	20.5	995	1	R84658	STLVpan-p pol prot
860	35	20.5	683	1	R11150	Fibulin C. Purifie	933	35	20.5	997	1	R66055	Human NMDA recepto
861	35	20.5	683	1	W27600	Human fibulin type	934	35	20.5	997	1	W85589	Human N-methyl-D-a
862	35	20.5	687	1	R62485	Human sVEGF-R1. Va	935	35	20.5	1055	1	R70114	IL-1rcypell-GPB 13
863	35	20.5	687	1	W47036	Soluble VEGF recep	936	35	20.5	1124	1	R41439	PT-NANBH NS5-NS3-c
864	35	20.5	687	1	W44841	Human soluble VEGF	937	35	20.5	1237	1	R13791	E75A protein. DNA
865	35	20.5	687	1	W68005	Human soluble vasc	938	35	20.5	1565	1	W18305	Photorehabdus lumin
866	35	20.5	703	1	W27601	Human fibulin type	939	35	20.5	3011	1	R34580	Human hepatitis C
867	35	20.5	708	1	R76236	Bovine foetal hear	940	35	20.5	3011	1	R34468	Encoded by full-le
868	35	20.5	709	1	W57329	Mitochondrial glyc	941	35	20.5	3038	1	R74171	Aspergillus terreu
869	35	20.5	709	1	W60262	Klebsiella pneumon	942	35	20.5	3079	1	R59926	GAP protein iraz2.
870	35	20.5	712	1	W30685	Glycerol kinase GU	943	35	20.5	3211	1	R81769	Sequence encoded b
871	35	20.5	712	1	W14306	Interleukin-1 rece	944	34.5	20.2	38	1	R27674	Endothelin antagon
872	35	20.5	718	1	W72245	HSV-2 strain SB5 C	945	34.5	20.2	87	1	R70736	PagD gene product.
873	35	20.5	728	1	W72246	HSV-2 strain SB5 C	946	34.5	20.2	87	1	W18381	S. typhimurium pag
874	35	20.5	734	1	W98740	H. pylori GHPO 961	947	34.5	20.2	106	1	R60160	Human spasmodytic
875	35	20.5	745	1	W17800	Myeloperoxidase. M	948	34.5	20.2	106	1	R60162	Human modified spa
876	35	20.5	754	1	R76226	Bovine endothelin	949	34.5	20.2	149	1	R96543	Hepatitis C virus
877	35	20.5	758	1	W13395	Amino-terminal fra	950	34.5	20.2	149	1	R96544	Hepatitis C virus
878	35	20.5	758	1	W36138	Human VEGF recepto	951	34.5	20.2	156	1	W88760	Polyptide fragne
879	35	20.5	758	1	W38369	Bovine endothelin	952	34.5	20.2	156	1	W88743	Secreted protein e
880	35	20.5	763	1	W31852	Mycobacterium tube	953	34.5	20.2	159	1	R60163	Human spasmodytic
881	35	20.5	771	1	W72247	HSV-2 strain SB5 C	954	34.5	20.2	180	1	R41108	Zea mays 15 kD see
882	35	20.5	777	1	R66044	Human N-methyl-D-a	955	34.5	20.2	192	1	R69676	Hepatitis C virus
883	35	20.5	777	1	W85578	Human N-methyl-D-a	956	34.5	20.2	192	1	R89547	Hepatitis C virus
884	35	20.5	780	1	R62487	Truncated Flt svEG	957	34.5	20.2	233	1	P94685	Amino acid sequenc
885	35	20.5	780	1	W47039	Soluble truncated	958	34.5	20.2	289	1	W19609	Pig growth hormone
886	35	20.5	780	1	W68008	Human soluble vasc	959	34.5	20.2	289	1	W19216	Swine growth hormo
887	35	20.5	805	1	W72248	HSV-2 strain SB5 C	960	34.5	20.2	314	1	Y01720	The MAGE-3 polyep
888	35	20.5	807	1	W06916	wSBE I-D4 amino ac	961	34.5	20.2	365	1	W51252	G-protein coupled
889	35	20.5	818	1	W72249	HSV-2 strain SB5 C	962	34.5	20.2	375	1	W51253	G-protein coupled
890	35	20.5	821	1	W72250	HSV-2 strain SB5 C	963	34.5	20.2	461	1	R07299	N-terminal region
891	35	20.5	842	1	W20894	H. pylori surface	964	34.5	20.2	470	1	W25766	Human CARR1. Genes
892	35	20.5	854	1	R66052	Human NMDA recepto	965	34.5	20.2	472	1	R76220	Spinach glycerol-3
893	35	20.5	854	1	W85586	Human N-methyl-D-a	966	34.5	20.2	479	1	P81036	Modified tissue pl
894	35	20.5	857	1	R19347	Human filamin-like	967	34.5	20.2	494	1	P70443	Sequence encoded b
895	35	20.5	870	1	R66051	Human NMDA recepto	968	34.5	20.2	512	1	P50629	Human parotid alph
896	35	20.5	870	1	W85585	Human N-methyl-D-a	969	34.5	20.2	512	1	P50630	Human pancreatic a
897	35	20.5	875	1	R66050	Human NMDA recepto	970	34.5	20.2	531	1	W51251	G-protein coupled
898	35	20.5	875	1	W85584	N-methyl-D-asparta	971	34.5	20.2	592	1	Y05369	Human HCMV Inducib
899	35	20.5	880	1	P80805	Sequence of env pr	972	34.5	20.2	594	1	W85552	Secreted protein e
900	35	20.5	885	1	R55530	Human NMDA R1e rec	973	34.5	20.2	595	1	W18097	Xenopus beta-signa
901	35	20.5	885	1	R57328	NDAR1-1. Human ce	974	34.5	20.2	604	1	W37163	Aspergillus oryzae
902	35	20.5	885	1	R80973	Human excitatory a	975	34.5	20.2	617	1	R42388	Consensus haemaggl
903	35	20.5	891	1	R66049	Human NMDA recepto	976	34.5	20.2	617	1	R42389	San Diego haemaggl
904	35	20.5	891	1	W85583	Human N-methyl-D-a	977	34.5	20.2	617	1	R42390	Chicago 1 haemaggl
905	35	20.5	908	1	R66053	Human NMDA recepto	978	34.5	20.2	617	1	R42391	Chicago 2 haemaggl
906	35	20.5	908	1	W85587	Human N-methyl-D-a	979	34.5	20.2	617	1	R42392	Mcl haemagglutinin
907	35	20.5	918	1	R66042	Human N-methyl-D-a	980	34.5	20.2	617	1	R49925	Consensus haemaggl
908	35	20.5	920	1	R49043	NMDA receptor chan	981	34.5	20.2	669	1	R86408	Human matrix metal
909	35	20.5	920	1	R49044	NMDA receptor chan	982	34.5	20.2	690	1	R04904	Residual protease-

CC thereby creating greater likelihood that the dimer will be formed in vivo
 CC and the formed dimer will have affinity for the native receptors and have
 CC agonistic activity on them. The changes stabilise the GPHs and prolong
 CC the biological activities of the hormones. The analogues can have uses
 CC as for the native GPHs.
 SQ Sequence 132 AA;

Query Match 95.3%; Score 163; DB 1; Length 132;
 Best Local Similarity 96.7%; Pred. No. 2.9e-15;
 Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcpplhvapyfsypva 30
 |||||
 Db 71 VCTYRDFIYRTVEIPQCPLHVAPYFSYPVA 100

RESULT 4
 R99419
 ID R99419 standard; Protein; 137 AA.
 AC R99419;
 DT 18-JAN-1997 (first entry)
 DE TSH beta subunit.
 KW Heteropolymetric protein; dimeric glycoprotein; hormone;
 KW thyroid stimulating hormone; TSH.
 OS Homo sapiens.
 PN EP-735139-AL.
 PD 02-OCT-1996.
 PF 20-JUN-1990; 306743.
 PR 20-JUN-1989; US-368628.
 PA (GENZ) GENZYME CORP.
 PI Chappel SC, Kelton CA, Nugent NP;
 DR WPI; 96-478414/48.
 DR N-PSDB; T41697.
 PT Increased recombinant protein production - esp. heteropolymetric
 PT glycoprotein hormones, uses DNA contg. at least one intron
 PT Example 3; Page 18-19; 26pp; English.
 PS The beta subunit (R99419) of human thyroid stimulating hormone
 CC (TSH) is encoded by an engineered genomic DNA fragment (T41697).
 CC Expression levels of recombinant TSH in mammalian (e.g. CHO) host
 CC cells are higher when cells are cotransfected with a vector
 CC including the beta subunit DNA and a vector including alpha
 CC subunit genomic DNA. When only alpha subunit cDNA is available,
 CC expression is improved by addition of an intron sequence into the
 CC cDNA construct.
 SQ Sequence 137 AA;

Query Match 91.5%; Score 156.5; DB 1; Length 137;
 Best Local Similarity 96.7%; Pred. No. 2.3e-14;
 Matches 29; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 vctyrdfiyrtveipgcpplhvapyfsypva 30
 |||||
 Db 71 VCTYRDFIYRT-EIPGCPPLHVAPYFSYPVA 99

RESULT 5
 R15092
 ID R15092 standard; Protein; 145 AA.
 AC R15092;
 DT 11-FEB-1992 (first entry)
 DE hCG/hTSH chimera, C5.
 KW Glycoprotein hormone; fertility; immuno-castration;
 KW immuno-contragative; vaccine; human chorionic gonadotropin;
 KW thyroid stimulating hormone; TSH; CG.
 OS Homo sapiens.
 PN WO9116922-A.
 PD 14-NOV-1991.
 PF 07-MAY-1991; U03162.
 PR 08-MAY-1990; US-520703.
 PA (UYNE-) UNIV MED NEW JERSEY.
 PI Campbell RK, Moyle WR;

DR WPI; 91-353528/48.
 PT New glyco-protein hormone analogues - for inducing fertility as
 PT immuno-castration agents, for suppressing reproductive system
 PT development and as immuno-contragative vaccines.
 PS Table III; Page 62; 94pp; English.
 CC The sequence is an analogue of mature hCG beta subunit having
 CC residues 74, 75, 77, 79, 80 and 83 replaced by the corresponding
 CC residues in the hTSH protein. The chimeric hormone may be useful
 CC as a TSH antagonist for the treatment of hyperthyroidism.
 CC See R15043, R15061-R15125 and R15161-R15198.
 SQ Sequence 145 AA;

Query Match 76.6%; Score 131; DB 1; Length 145;
 Best Local Similarity 70.0%; Pred. No. 7.2e-11;
 Matches 21; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcpplhvapyfsypva 30
 |||||
 Db 56 VCTYRDFIYRTVEIPGCPPLHVAPYFSYPVA 85

RESULT 6
 R15091
 ID R15091 standard; Protein; 145 AA.
 AC R15091;
 DT 11-FEB-1992 (first entry)
 DE hCG/hTSH chimera, C4.
 KW Glycoprotein hormone; fertility; immuno-castration;
 KW immuno-contragative; vaccine; human chorionic gonadotropin;
 KW thyroid stimulating hormone; TSH; CG.
 OS Homo sapiens.
 PN WO9116922-A.
 PD 14-NOV-1991.
 PF 07-MAY-1991; U03162.
 PR 08-MAY-1990; US-520703.
 PA (UYNE-) UNIV MED NEW JERSEY.
 PI Campbell RK, Moyle WR;
 DR WPI; 91-353528/48.
 PT New glyco-protein hormone analogues - for inducing fertility as
 PT immuno-castration agents, for suppressing reproductive system
 PT development and as immuno-contragative vaccines.
 PS Table III; Page 62; 94pp; English.
 CC The sequence is an analogue of mature hCG beta subunit having
 CC residues 62-69 replaced by the corresponding residues in the hTSH
 CC protein. The chimeric hormone may be useful as a TSH antagonist
 CC for the treatment of hyperthyroidism.
 CC See R15043, R15061-R15125 and R15161-R15198.
 SQ Sequence 145 AA;

Query Match 70.8%; Score 121; DB 1; Length 145;
 Best Local Similarity 76.7%; Pred. No. 1.7e-09;
 Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcpplhvapyfsypva 30
 |||||
 Db 56 VCTYRDFIYRTVEIPGCPRGVNPVSYAVA 85

RESULT 7
 W27683
 ID W27683 standard; protein; 145 AA.
 DT 12-JAN-1998 (first entry)
 DE Chorionic gonadotropin beta subunit carboxy-terminal loop mutant.
 KW Human; chorionic gonadotropin; chorionic gonadotropin; beta-hCG;
 KW beta subunit; carboxy-terminal loop; mutant; reduction; vaccine;
 KW contragative medicament; cross-reactivity; luteinising hormone;
 KW contraceptive; immunoassay; Kaposi sarcoma; inhibition; LH;
 KW neutralising antibody.
 OS Homo sapiens.
 OS Synthetic.


```

FH Key                               Location/Qualifiers
FT peptide                          1..5
FT protein                          1..140
PN W08607383-A.
PD 18-DEC-1986.
PF 04-JUN-1986; U01226.
PR 04-JUN-1985; US-741168.
PR 18-JUL-1985; US-756847.
PA (BIOT-) BIOTECHN RES PARTN.
PI Taimadge KD, Fiddes JC;
DR WPI; 86-346608/52.
DR N-PSDB; N60524.
PT Auto-antigen vaccines conferring antigenicity using multimers
PT etc. - useful as species specific or cross-species effective,
PT esp. for controlling fertility in mammals
PS Example; Fig 4; 101pp; English.
CC The patentors claim a vaccine effective against mammalian fertility
CC comprising a vaccinia virus genome having disposed unit, in a non-
CC essential region, a DNA sequence of formula : (Hormone)n; n= 1-20;
CC Hormone= DNA sequence derived from the sequence encoding a
CC reproductive hormone. The hormone is esp. LH, GnRH, CG or FSH.
SQ Sequence 141 AA;

Query Match 55.6%; Score 95; DB 1; Length 141;
Best Local Similarity 53.3%; Pred. No. 5.5e-06;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfrtveipgcplhvapyfsypva 30
   ||||| : : : ||| | : |||
Db 76 VCTYRDVFESIRLPGCRGVDPVVSPVA 105

RESULT 12
W99547
ID W99547 standard; Protein; 141 AA.
AC W99547;
DT 08-JUN-1999 (first entry)
DE hLH-beta analogue hLH-beta-Y37C.
KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
KW human follicle stimulating hormone; human thyroid stimulating hormone.
KW stability; primer; amplification; PCR; mutation.
OS Homo sapiens.
OS Synthetic.
PN W09858957-A2.
PD 30-DEC-1998.
PF 25-JUN-1998; U13070.
PR 25-JUN-1997; US-050784.
PA (ISTF ) ARS APPLIED RES SYSTEMS HOLDING NV.
PA (MCIN-) MCINNIS P G.
PI Moyle WR;
DR WPI; 99-081219/07.
PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
PT or hTSH, have an intersubunit disulphide crosslink between the
PT alpha- and beta-subunits to improve stability
PS Disclosure; Fig 36; 139pp; English.
CC The invention relates to the production of analogues of a heterodimeric
CC subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
CC (hCG), human luteinising hormone (hLH), human follicle stimulating
CC hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
CC human chorionic gonadotropin; human luteinising hormone; disulphide bond;
CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
CC improved stability, the analogue retaining at least a portion of the
CC bioactivity for the corresponding native GPH receptor. This sequence
CC represents a mutant hLH-beta subunit used for the generation of
CC the modified GPHs. The improved analogues are designed specifically
CC to reduce perturbation of the 3-dimensional structure of the hormone,
CC thereby creating greater likelihood that the dimer will be formed in vivo
CC and the formed dimer will have affinity for the native receptors and have
CC agonistic activity on them. The changes stabilise the GPHs and prolong
CC the biological activities of the hormones. The analogues can have uses
CC as for the native GPHs.
SQ Sequence 141 AA;

Query Match 55.6%; Score 95; DB 1; Length 141;
Best Local Similarity 53.3%; Pred. No. 5.5e-06;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfrtveipgcplhvapyfsypva 30
   ||||| : : : ||| | : |||
Db 76 VCTYRDVFESIRLPGCRGVDPVVSPVA 105

RESULT 13
W99520
ID W99520 standard; Protein; 141 AA.
AC W99520;
DT 08-JUN-1999 (first entry)
DE Glycoprotein hormone analogue hLH-beta-133C.
KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
KW human follicle stimulating hormone; human thyroid stimulating hormone;
KW as for the native GPHs.

```



```

RESULT 16
W99523
ID W99523 standard; Protein; 141 AA.
AC W99523;
DT 08-JUN-1999 (first entry)
DE Glycoprotein hormone analogue hLH-beta-M42C.
KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
KW human follicle stimulating hormone; human thyroid stimulating hormone;
KW stability; primer; amplification; PCR; mutation.
OS Homo sapiens.
OS Synthetic.
PN W09858957-A2.
PD 30-DEC-1998.
PF 25-JUN-1998; U13070.
PR (ISTF ) ARS APPLIED RES SYSTEMS HOLDING NV.
PA (MCIN-) MCINNIS P G.
PI Moyle WR;
DR WPI; 99-081219/07.

PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
PT or hTSH, have an intersubunit disulphide crosslink between the
PT alpha- and beta-subunits to improve stability
PS Example 12; Page 90; 139pp; English.
CC The invention relates to the production of analogues of a heterodimeric
CC subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
CC (hCG), human luteinising hormone (hLH), human follicle stimulating
CC hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
CC muteins, which are modified to contain an intersubunit disulphide bond,
CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
CC improved stability, the analogue retaining at least a portion of the
CC bioactivity for the corresponding native GPH receptor. This sequence
CC represents a mutated hLH-beta subunit used for the generation of the
CC modified GPHs. The improved analogues are designed specifically to
CC reduce perturbation of the 3-dimensional structure of the hormone,
CC thereby creating greater likelihood that the dimer will be formed in vivo
CC and the formed dimer will have affinity for the native receptors and have
CC agonistic activity on them. The changes stabilise the GPHs and prolong
CC the biological activities of the hormones. The analogues can have uses
CC as for the native GPHs.
SQ Sequence 141 AA;

Query Match 55.6%; Score 95; DB 1; Length 141;
Best Local Similarity 53.3%; Pred. No. 5.5e-06;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpvhvafsyypva 30
||||| : : : |||| | | : |||
DB 76 VCTYRDVFESIRLPGCGVDPVWSFPVA 105

RESULT 17
W99524
ID W99524 standard; Protein; 141 AA.
AC W99524;
DT 08-JUN-1999 (first entry)
DE Glycoprotein hormone analogue hLH-beta-Q46C.
KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
KW human follicle stimulating hormone; human thyroid stimulating hormone;
KW stability; primer; amplification; PCR; mutation.
OS Homo sapiens.
OS Synthetic.
PN W09858957-A2.
PD 30-DEC-1998.
PF 25-JUN-1998; U13070.
PR (ISTF ) ARS APPLIED RES SYSTEMS HOLDING NV.
PA (MCIN-) MCINNIS P G.
PI Moyle WR;
DR WPI; 99-081219/07.

PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
PT or hTSH, have an intersubunit disulphide crosslink between the
PT alpha- and beta-subunits to improve stability
PS Example 12; Page 90; 139pp; English.
CC The invention relates to the production of analogues of a heterodimeric
CC subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
CC (hCG), human luteinising hormone (hLH), human follicle stimulating
CC hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
CC muteins, which are modified to contain an intersubunit disulphide bond,
CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
CC improved stability, the analogue retaining at least a portion of the
CC bioactivity for the corresponding native GPH receptor. This sequence
CC represents a mutated hLH-beta subunit used for the generation of the
CC modified GPHs. The improved analogues are designed specifically to
CC reduce perturbation of the 3-dimensional structure of the hormone,
CC thereby creating greater likelihood that the dimer will be formed in vivo
CC and the formed dimer will have affinity for the native receptors and have
CC agonistic activity on them. The changes stabilise the GPHs and prolong
CC the biological activities of the hormones. The analogues can have uses
CC as for the native GPHs.
SQ Sequence 141 AA;

Query Match 55.6%; Score 95; DB 1; Length 141;
Best Local Similarity 53.3%; Pred. No. 5.5e-06;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpvhvafsyypva 30
||||| : : : |||| | | : |||
DB 76 VCTYRDVFESIRLPGCGVDPVWSFPVA 105

```

```

PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
PT or hTSH, have an intersubunit disulphide crosslink between the
PT alpha- and beta-subunits to improve stability
PS Example 12; Page 90; 139pp; English.
CC The invention relates to the production of analogues of a heterodimeric
CC subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
CC (hCG), human luteinising hormone (hLH), human follicle stimulating
CC hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
CC muteins, which are modified to contain an intersubunit disulphide bond,
CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
CC improved stability, the analogue retaining at least a portion of the
CC bioactivity for the corresponding native GPH receptor. This sequence
CC represents a mutated hLH-beta subunit used for the generation of the
CC modified GPHs. The improved analogues are designed specifically to
CC reduce perturbation of the 3-dimensional structure of the hormone,
CC thereby creating greater likelihood that the dimer will be formed in vivo
CC and the formed dimer will have affinity for the native receptors and have
CC agonistic activity on them. The changes stabilise the GPHs and prolong
CC the biological activities of the hormones. The analogues can have uses
CC as for the native GPHs.
SQ Sequence 141 AA;

Query Match 55.6%; Score 95; DB 1; Length 141;
Best Local Similarity 53.3%; Pred. No. 5.5e-06;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpvhvafsyypva 30
||||| : : : |||| | | : |||
DB 76 VCTYRDVFESIRLPGCGVDPVWSFPVA 105

RESULT 18
W99525
ID W99525 standard; Protein; 141 AA.
AC W99525;
DT 08-JUN-1999 (first entry)
DE Glycoprotein hormone analogue hLH-beta-Q54C.
KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
KW human follicle stimulating hormone; human thyroid stimulating hormone;
KW stability; primer; amplification; PCR; mutation.
OS Homo sapiens.
OS Synthetic.
PN W09858957-A2.
PD 30-DEC-1998.
PF 25-JUN-1998; U13070.
PR (ISTF ) ARS APPLIED RES SYSTEMS HOLDING NV.
PA (MCIN-) MCINNIS P G.
PI Moyle WR;
DR WPI; 99-081219/07.

PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
PT or hTSH, have an intersubunit disulphide crosslink between the
PT alpha- and beta-subunits to improve stability
PS Example 12; Page 90; 139pp; English.
CC The invention relates to the production of analogues of a heterodimeric
CC subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
CC (hCG), human luteinising hormone (hLH), human follicle stimulating
CC hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
CC muteins, which are modified to contain an intersubunit disulphide bond,
CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
CC improved stability, the analogue retaining at least a portion of the
CC bioactivity for the corresponding native GPH receptor. This sequence
CC represents a mutated hLH-beta subunit used for the generation of the
CC modified GPHs. The improved analogues are designed specifically to
CC reduce perturbation of the 3-dimensional structure of the hormone,
CC thereby creating greater likelihood that the dimer will be formed in vivo
CC and the formed dimer will have affinity for the native receptors and have
CC agonistic activity on them. The changes stabilise the GPHs and prolong
CC the biological activities of the hormones. The analogues can have uses
CC as for the native GPHs.
SQ Sequence 141 AA;

```


Query Match 55.6%; Score 95; DB 1; Length 141;
 Best Local Similarity 53.3%; Pred. No. 5.5e-06;
 Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcpplhvapyfsypva 30
 ||||| : : : ||||| : : |||||
 Db 76 VCTYRDVFESIRLPGCPRGVDPVVSFPVA 105

RESULT 19
 W99516
 ID W99516 standard; Protein; 141 AA.
 AC W99516;
 DT 08-JUN-1999 (first entry)
 DE Glycoprotein hormone analogue hLH-beta-V44C.
 KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
 KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
 KW human follicle stimulating hormone; human thyroid stimulating hormone;
 KW stability; primer; amplification; PCR; mutation.
 OS Homo sapiens.
 PN WO9858957-A2.
 PD 30-DEC-1998.
 PF 25-JUN-1998; U13070.
 PR 25-JUN-1997; US-050784.
 PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
 PA (MCIN-) MCINNIS P G.
 PI Moyle WR;
 DT WPI; 99-081219/07.
 PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
 PT or hTSH, have an intersubunit disulphide crosslink between the
 PT alpha- and beta-subunits to improve stability
 PS Example 12; Page 90; 139pp; English.
 CC The invention relates to the production of analogues of a heterodimeric
 CC subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
 CC (hCG), human luteinising hormone (hLH), human follicle stimulating
 CC hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
 CC muteins, which are modified to contain an intersubunit disulphide bond,
 CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
 CC improved stability, the analogue retaining at least a portion of the
 CC bioactivity for the corresponding native GPH receptor. This sequence
 CC represents a mutated hLH-beta subunit used for the generation of the
 CC modified GPHs. The improved analogues are designed specifically to
 CC reduce perturbation of the 3-dimensional structure of the hormone,
 CC thereby creating greater likelihood that the dimer will be formed in vivo
 CC and the formed dimer will have affinity for the native receptors and have
 CC agonistic activity on them. The changes stabilise the GPHs and prolong
 CC the biological activities of the hormones. The analogues can have uses
 CC as for the native GPHs.
 SQ Sequence 141 AA;

Query Match 55.6%; Score 95; DB 1; Length 141;
 Best Local Similarity 53.3%; Pred. No. 5.5e-06;
 Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcpplhvapyfsypva 30
 ||||| : : : ||||| : : |||||
 Db 76 VCTYRDVFESIRLPGCPRGVDPVVSFPVA 105

RESULT 20
 W99517
 ID W99517 standard; Protein; 141 AA.
 AC W99517;
 DT 08-JUN-1999 (first entry)
 DE Glycoprotein hormone analogue hLH-beta-M41C.
 KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
 KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
 KW human follicle stimulating hormone; human thyroid stimulating hormone;
 KW stability; primer; amplification; PCR; mutation.

OS Homo sapiens.
 OS Synthetic.
 PN WO9858957-A2.
 PD 30-DEC-1998.
 PF 25-JUN-1998; U13070.
 PR 25-JUN-1997; US-050784.
 PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
 PA (MCIN-) MCINNIS P G.
 PI Moyle WR;
 DT WPI; 99-081219/07.
 PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
 PT or hTSH, have an intersubunit disulphide crosslink between the
 PT alpha- and beta-subunits to improve stability
 PS Example 12; Page 90; 139pp; English.
 CC The invention relates to the production of analogues of a heterodimeric
 CC subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
 CC (hCG), human luteinising hormone (hLH), human follicle stimulating
 CC hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
 CC muteins, which are modified to contain an intersubunit disulphide bond,
 CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
 CC improved stability, the analogue retaining at least a portion of the
 CC bioactivity for the corresponding native GPH receptor. This sequence
 CC represents a mutated hLH-beta subunit used for the generation of the
 CC modified GPHs. The improved analogues are designed specifically to
 CC reduce perturbation of the 3-dimensional structure of the hormone,
 CC thereby creating greater likelihood that the dimer will be formed in vivo
 CC and the formed dimer will have affinity for the native receptors and have
 CC agonistic activity on them. The changes stabilise the GPHs and prolong
 CC the biological activities of the hormones. The analogues can have uses
 CC as for the native GPHs.
 SQ Sequence 141 AA;

Query Match 55.6%; Score 95; DB 1; Length 141;
 Best Local Similarity 53.3%; Pred. No. 5.5e-06;
 Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcpplhvapyfsypva 30
 ||||| : : : ||||| : : |||||
 Db 76 VCTYRDVFESIRLPGCPRGVDPVVSFPVA 105

RESULT 21
 W99519
 ID W99519 standard; Protein; 141 AA.
 AC W99519;
 DT 08-JUN-1999 (first entry)
 DE Glycoprotein hormone analogue hLH-beta-A35C.
 KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
 KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
 KW human follicle stimulating hormone; human thyroid stimulating hormone;
 KW stability; primer; amplification; PCR; mutation.
 OS Homo sapiens.
 OS Synthetic.
 PN WO9858957-A2.
 PD 30-DEC-1998.
 PF 25-JUN-1998; U13070.
 PR 25-JUN-1997; US-050784.
 PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
 PA (MCIN-) MCINNIS P G.
 PI Moyle WR;
 DT WPI; 99-081219/07.
 PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
 PT or hTSH, have an intersubunit disulphide crosslink between the
 PT alpha- and beta-subunits to improve stability
 PS Example 12; Page 90; 139pp; English.
 CC The invention relates to the production of analogues of a heterodimeric
 CC subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
 CC (hCG), human luteinising hormone (hLH), human follicle stimulating
 CC hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
 CC muteins, which are modified to contain an intersubunit disulphide bond,
 CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
 CC improved stability, the analogue retaining at least a portion of the

bioactivity for the corresponding native GPH receptor. This sequence represents a mutated hGH-beta subunit used for the generation of the modified GPHs. The improved analogues are designed specifically to reduce perturbation of the 3-dimensional structure of the hormone, thereby creating greater likelihood that the dimer will be formed in vivo and the formed dimer will have affinity for the native receptors and have agonistic activity on them. The changes stabilise the GPHs and prolong the biological activities of the hormones. The analogues can have uses as for the native GPHs.

Sequence 141 AA;

Query Match 55.6%; Score 95; DB 1; Length 141;
Best Local Similarity 53.3%; Pred. NO. 5.5e-06;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcpvhvapyfypva 30
 ||||| : :: ||||| | | : |||||
 Db 76 VCTYRDVFEEISRLPGCPRGVDFVVSFPA 105

RESULT	22	
R86261		
ID	R86261 standard; Protein; 234 AA.	
AC	R86261;	
DT	08-MAY-1996 (first entry)	
DE	Partially deglycosylated single chain gonadotropin analogue 3a.	
KW	Single chain gonadotropin; human chorionic gonadotropin; hCG;	
KW	alpha; beta; subunit; analogue; glycoprotein hormone; fertility;	
KW	inhibit; stimulate; increase; lutropin; luteinising hormone; LH;	
KW	folicle stimulating hormone; FSH; vaccine; contraceptive.	
OS	Synthetic.	
FH	Key	Location/Qualifiers
FT	peptide	1..20
FT		/label= leader
FT	region	21..134
FT		/label= h_uH_beta_subunit_(1-114)
FT	misc_difference 50	
FT		/note= "wild-type Asn at position 30 of the beta-
FT		subunit is pref. replaced by another amino
FT		acid (esp. Gln) to remove a glycosylation
FT		site"
FT	region	135..142
FT		/label= linker
FT	region	143..234
FT		/label= Gonadotropin_alpha_subunit_(1-92)
FT	misc_difference 194	
FT		/note= "wild-type Asn at position 52 of the alpha
FT		subunit is pref. replaced by another amino
FT		acid (esp. Gln) to remove a glycosylation
FT		site"
FT	misc_difference 220	
FT		/note= "wild-type Asn at position 78 of the alpha
FT		subunit is pref. replaced by another amino
FT		acid (esp. Gln) to remove a glycosylation
FT		site"
FT		

PN WO9522340-A1.
PD 24-AUG-1995.
PF 17-FEB-1995; U02067.
PF 18-FEB-1994; U5199382.
PR (SENS-) SENS1-TEST.
PI Moyle WR:
PR WPI; 95-302553/39.
PT Methods for altering fertility in mammals, esp. humans - e.g.
PT stimulating fertility by reducing the activity and/or levels of
PT circulating glyco:protein hormones having lutropin activity
PS Example 24; Fig 8 and Page 60; 102pp; English.
PS The single-chain gonadotropin analogue 3a (human LH-beta(1-114)
CC [N30X]-linker-human CG-alpha(1-92)[N52X,N78X]) can be derived
CC from analogue 3 by removing at least one of the glycosylation
CC sites. The partially deglycosylated analogue has anti-luteinising
CC hormone (lutropin) activity and can be used for facilitating
CC ovulation, terminating pregnancy and reducing androgen secretion.

SQ Sequence 234 AA;

Query Match 55.6%; Score 95; DB 1; Length 234;
Best Local Similarity 53.3%; Pred. No. 9.1e-06;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfflyrtvdlpgcplhvapyfsypva 30
||||| : :: : ||||| | | : |||
Db 76 VCTYRDVFESRLPGCERGVDPVVSFPA 105

```

RESULT 23
R86249 ID R86249 standard; Protein; 234 AA.
R86249 AC
R86249 DT 26-APR-1996 (first entry)
R86249 DE Single chain gonadotropin analogue 3.
R86249 KW Single chain gonadotropin; human chorionic gonadotropin; hCG;
R86249 KW alpha; beta; subunit; analogue; glycoprotein hormone; fertility;
R86249 KW inhibit; stimulate; increase; lutropin; luteinising hormone; LH;
R86249 KW follicle stimulating hormone; FSH; vaccine; contraceptive.
R86249 OS Synthetic.
R86249 FH key Location/Qualifiers
R86249 FH peptide 1..20
R86249 FT FT /label= leader
R86249 FT FT 21..134
R86249 FT FT /label= hLH_beta_subunit_(1-114)
R86249 FT FT 135..142
R86249 FT FT /label= linker
R86249 FT FT 143..234
R86249 FT FT /label= Gonadotropin_alpha_subunit_(1-92)
R86249 PN W09523340-A1.
R86249 PD 24-AUG-1995.
R86249 PF 17-FEB-1995; U02067.
R86249 PR 18-FEB-1994; US-199382.
R86249 PA (SENS-) SENS-TEST.
R86249 PI MOyle WR.
R86249 DR WFI; 95-302553/39.
R86249 DR N-PSDB; T03221.
R86249 PT Methods for altering fertility in mammals, esp. humans - e.g.
R86249 PT stimulating fertility by reducing the activity and/or levels of
R86249 PT circulating glyco:protein hormones having lutropin activity
R86249 PS Example 14; Fig 8; 102pp; English.
R86249 CC Analogue 3 (human LH-beta(1-114))-linker-human-alpha(1-92)) is a
R86249 CC specific example of a single chain gonadotropin; chimeric protein
R86249 CC having a chorionic gonadotropin (CG) beta-subunit at the N-termi
R86249 CC and a CG alpha-subunit at the C-terminus, joined by a linker of
R86249 CC amino acids are claimed. The analogue has luteinising hormone
R86249 CC (lutropin) activity and is useful for inducing ovulation and
R86249 CC increasing male fertility.
R86249 S Sequence 234 AA.

```

Query Match 55.6%; Score 95; DB 1; Length 234;
Best Local Similarity 53.3%; Pred. No. 9.1e-06;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Caps 0;

Qy 1 vctyrdnlyrtvstpgcdhvpayfsypva 30
||||| : :: ||||| | | : |||||
ph 76 vctyrdnlyrtvstpgcdhvpayfsypva 105

RESULT	24
R10038	
ID	R10038 standard; Protein; 129 AA.
AC	R10038;
DT	10-APR-1991 (first entry)
DE	Ovine FSH beta subunit precursor encoded by clone 5.14.
KE	Follicle stimulating hormone; ovulation.
KW	Ovis ammon aries.
FT	key
FT	Location/Qualifiers
FT	peptide 1..19

CC by exons II and III. This sequence was used in the production of an
 CC extended FSH beta subunit in which the carboxy terminal peptide, ie.
 CC residues 112-118 to 145, of human chorionic gonadotropin beta
 CC subunit to position 111 of wild type FSH beta subunit. The modified
 CC form of the FSH beta subunit enhances the rate of dimerisation and
 CC secretion of dimers and individual chains. An FSH alpha-beta hetero-
 CC dimer is useful in a pharmaceutical composition for therapeutic use
 CC in the regulation of metabolism pertinent to reproduction in human
 CC females.
 CC Sequence 76 AA;

Query Match 53.2%; Score 91; DB 1; Length 76;
 Best Local Similarity 44.8%; Pred. No. 1e-05; Mismatches 9; Indels 0; Gaps 0;
 Matches 13; Conservative 7;

QY 2 ctgydfiyrvtveipgcpplhvafysypva 30
 ||::: || || || | :|||
 Db 16 CTFKELVYETVRVPGCAHADSLSLYTPVA 44

RESULT 28
 R73939
 ID R73939 standard; Protein; 76 AA.
 AC R73939;
 DT 23-OCT-1995 (first entry)
 DE Human FSH-beta exon III product.
 KW FSH-beta; follicle stimulating hormone; glycosylation.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Protein 1..76
 FT /label = Mat_peptide
 FT /note= "amino acids 36-111 of mature FSH-beta"
 FT
 PD US5405945-A.
 PN 11-APR-1995.
 PE 21-FEB-1989; 313646.
 PR 21-FEB-1989; US-513646.
 PR 01-JUN-1990; US-532254.
 PR 24-SEP-1992; US-950835.
 PA (UNIW) UNIV WASHINGTON.
 PI Boine I, Matzuk MM;
 DR WPI: 95-154629/20.
 DR N-PSDB; Q91983.
 PT DNA encoding human gonadotropin alpha subunit - used with beta
 PT subunit DNA for the recombinant prodn of hCG, FSH, LH or TSH
 PS Disclosure: Fig.2A-B; 28pp; English.
 CC A probe based on amino acids 94-107 of exon III of human FSH-beta
 CC was used to isolate an FSH-beta gene (Q91983) from a library of
 CC human placental JAR choriocarcinoma genomic DNA in lambda MG3.
 CC Exon II of the gene encoded the signal peptide and amino acids
 CC 1-35 of the mature FSH-beta.
 CC Sequence 76 AA;

Query Match 53.2%; Score 91; DB 1; Length 76;
 Best Local Similarity 44.8%; Pred. No. 1e-05;
 Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 2 ctgydfiyrvtveipgcpplhvafysypva 30
 ||::: || || || | :|||
 Db 16 CTFKELVYETVRVPGCAHADSLSLYTPVA 44

RESULT 29
 P92145
 ID P92145 standard; protein; 110 AA.
 AC P92145;
 DT 26-JUN-1990 (first entry)
 DE Beta subunit of urinary follicle stimulating hormone.
 KW Urinary follicle stimulating hormone; beta subunit;
 KW human menopausal gonadotropin; infertility.
 OS Homo sapiens.
 PN W08810270-A.

PD 29-DEC-1988.
 PF 24-JUN-1987; IT0048.
 PR 26-JUN-1987; IT-048110.
 PA (ICE-) Ist Ric Cesare Sero.
 PI Arpaia G, Serani S, Villa S;
 DR WPI: 89-023821/03.
 PT Urinary follicle stimulating hormone - is biologically active and free
 PT from detectable traces of luteinising hormone and other urinary hormones.
 PS Claim 1; page 21; 32pp; English.
 CC Residues 1(Asn) and 2(Ser) are optional. The peptide has 111 amino acids
 CC instead of 118 or 108 as reported for the known FSH beta subunit. It was
 CC purified from human menopausal gonadotropin. Together with an alpha
 CC subunit (sequence known from J Biol. Chem. 250, 6735 (1975)) the
 CC protein forms an FSH protein which can be used to treat infertility.
 CC Sequence 110 AA;

Query Match 53.2%; Score 91; DB 1; Length 110;
 Best Local Similarity 44.8%; Pred. No. 1.5e-05;
 Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 2 ctgydfiyrvtveipgcpplhvafysypva 30
 ||::: || || || | :|||
 Db 51 CTFKELVYETVRVPGCAHADSLSLYTPVA 79

RESULT 30
 R31000
 ID R31000 standard; protein; 111 AA.
 AC R31000;
 DT 14-MAY-1993 (first entry)
 DE Human follicle stimulating hormone beta-subunit.
 KW hCG; glycoprotein hormone analogue; human infertility; LH; FSH;
 KW luteinising hormone receptor; follicle stimulating hormone receptor;
 KW vertebrate; polycystic ovarian disease.
 OS Homo sapiens.
 PN W09222588-A.
 PD 23-DEC-1992.
 PF 18-JUN-1991; U05207.
 PR 18-JUN-1991; US-717151.
 PA (UYNE-) UNIV NEW JERSEY.
 PI Campbell RK, Moyle WR,
 DR WPI: 93-018070/02.
 PT New alpha, beta-hetero:dimeric polypeptide derivs. - which bind to
 PT luteinising and follicle stimulating hormone receptors, useful for
 PT controlling the ratio of FSH to LH activity
 PS Disclosure: Page 20; 98pp; English.
 CC The sequence is that of the beta-subunit of human follicle stimulating
 CC hormone which was used in the creation of a beta-subunit peptide
 CC analogue as part of an alpha, beta-heterodimeric polypeptide
 CC having an affinity to vertebrate luteinising hormone (LH) receptors
 CC and vertebrate follicle stimulating hormone (FSH) receptors. This can
 CC be used for treating human infertility or polycystic ovarian disease.
 CC Sequence 111 AA;

Query Match 53.2%; Score 91; DB 1; Length 111;
 Best Local Similarity 44.8%; Pred. No. 1.5e-05;
 Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 2 ctgydfiyrvtveipgcpplhvafysypva 30
 ||::: || || || | :|||
 Db 51 CTFKELVYETVRVPGCAHADSLSLYTPVA 79

RESULT 31
 R15086
 ID R15086 standard; Protein; 112 AA.
 AC R15086;
 DT 11-FEB-1992 (first entry)
 DE hCG/hFSH chimera, B26.
 KW Glycoprotein hormone; fertility; immuno-castration;
 KW immuno-contragestive; vaccine; human chorionic gonadotropin;


```

KW follicle stimulating hormone; FSH; CG.
OS Homo sapiens.
PN W09116922-A.
PD 14-NOV-1991.
PF 07-MAY-1991; U03162.
PR 08-MAY-1990; US-520703.
PA (UYNE-) UNIV MED NEW JERSEY.
PI Campbell RK, Moyle WR;
DR WPI: 91-353528/48.
PT New glyco-protein hormone analogues - for inducing fertility as
PT immuno-castration agents, for suppressing reproductive system
PT development and as immuno-contragestive vaccines.
PS Table II; Page 61; 94pp; English.
CC The sequence is an analogue of mature hFSH beta subunit having
CC several residues replaced by the corresponding residues in the hCG
CC protein. The chimeric hormone is capable of directing hormone
CC binding to both LH and FSH receptors and may be useful for the
CC treatment of infertility in men and women and the promotion of
CC fertility in male and female animals. (See R15043, R15061-R15125 and
CC R15161-R15198).
SQ Sequence 112 AA;

Query Match 53.2%; Score 91; DB 1; Length 112;
Best Local Similarity 44.8%; Pred. No. 1.5e-05;
Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 2 cttyrdfflyrtveipgcpvhvafsyypva 30
Db 51 CTFKELVYETVRVPGCAHHAHDSLYTPVA 79

RESULT 32
R15082
ID R15082 standard; Protein; 118 AA.
AC R15082;
DT 11-FEB-1992 (first entry)
DE hCG/hFSH chimera, B22.
KW Glycoprotein hormone; fertility; immuno-castration;
KW immuno-contragestive; vaccine; human chorionic gonadotropin;
KW follicle stimulating hormone; FSH; CG.
OS Homo sapiens.
PN W09116922-A.
PD 14-NOV-1991.
PF 07-MAY-1991; U03162.
PR 08-MAY-1990; US-520703.
PA (UYNE-) UNIV MED NEW JERSEY.
PI Campbell RK, Moyle WR;
DR WPI: 91-353528/48.
PT New glyco-protein hormone analogues - for inducing fertility as
PT immuno-castration agents, for suppressing reproductive system
PT development and as immuno-contragestive vaccines.
PS Table II; Page 61; 94pp; English.
CC The sequence is an analogue of mature hFSH beta subunit having
CC several residues replaced by the corresponding residues in the hCG
CC protein. The chimeric hormone is capable of directing hormone
CC binding to both LH and FSH receptors and may be useful for the
CC treatment of infertility in men and women and the promotion of
CC fertility in male and female animals. (See R15043, R15061-R15125 and
CC R15161-R15198).
SQ Sequence 118 AA;

Query Match 53.2%; Score 91; DB 1; Length 118;
Best Local Similarity 44.8%; Pred. No. 1.6e-05;
Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 2 cttyrdfflyrtveipgcpvhvafsyypva 30
Db 57 CTFKELVYETVRVPGCAHHAHDSLYTPVA 85

RESULT 33

```

```

R06724
ID R06724 standard; protein; 129 AA.
AC R06724;
DT 23-JAN-1991 (first entry)
DE Human follicle-stimulating hormone beta.
KW Human follicle-stimulating hormone-beta; FSH; modification; metabolism;
KW reproductive disorder.
OS Homo sapiens.
PH Key Location/Qualifiers
FT peptide 1..18
FT /label=signal_peptide
FT protein 18..0
FT /label=mature_protein
PN W09009800-A.
PD 07-SEP-1990.
PF 20-FEB-1990; U01037.
PR 21-FEB-1989; US-313646.
PA (UNIW) UNIV OF WASHINGTON.
PI Bolme I, Matzuk MM, Keene JL;
DR WPI: 90-250155/38.
DR N-FSDB; Q06974.
PT Modified forms of fsh, lh, cg and tsh - efficiently produced and
PT secreted by recombinant techniques with altered glycosylation and
PT activity
PS Disclosure: Fig 2A-B; 68pp; English.
CC The amino acid sequence is identical to that reported by Watkins,
CC T.C. et al. DNA (1987) 6:205-212. But the carboxy terminal
CC sequence IPTALSY reported by Saxena, D.B., J Biol Chem (1976) 251:
CC 993-1005 is found neither in the sequence represented here, nor in
CC the protein based sequence reported by Shome, B., et al., J Clin
CC Endocrinol Metab (1974) 39:203-205. A more recent determination of
CC the amino acid sequence confirms the sequence deduced from the DNA
CC (Stone, B. et al., J Prot Chem (1988) 7:325-339.
CC An important modification of the beta-FSH chain encoding DNA is
CC obtained when the 34 amino acid carboxy terminal peptide of the
CC chorionic gonadotropin beta chain is fused to the C-terminus.
CC In this form of the hormone, the C-terminal Glu of FSH beta at
CC position 111 is extended with the amino acid sequence representing from
CC about amino acids 112-118 to 145 of the beta CG sequence.
CC The recombinant hormone can be produced with altered glycosylation
CC or activity patterns. It can be used to regulate reproductive
CC metabolism and to treat reproductive disorders, and has prolonged
CC circulating half-life in the body.
SQ Sequence 129 AA;

Query Match 53.2%; Score 91; DB 1; Length 129;
Best Local Similarity 44.8%; Pred. No. 1.8e-05;
Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 2 cttyrdfflyrtveipgcpvhvafsyypva 30
Db 69 CTFKELVYETVRVPGCAHHAHDSLYTPVA 97

RESULT 34
R10096
ID R10096 standard; Protein; 129 AA.
AC R10096;
DT 14-MAR-1991 (first entry)
DE Engineered partial human FSH beta subunit.
KW human follicle stimulating hormone subunit beta; FSH;
KW heteropolymeric protein production.
OS Synthetic.
PH Key Location/Qualifiers
FT region 1..53
FT /note= "encoded by exon II"
FT region 54..129
FT /note= "encoded by exon III"
PN EP-404546-A.
PD 27-DEC-1990.
PF 20-JUN-1990; 306743.
PR 20-JUN-1989; US-368628.

```



```

PA (GENZ-) GENZYME CORP.
PI Kelton CA, Nugent NP, Cheppel SC;
DR WPI: 91-001567/01.
DR Q-PSDB: Q10076, Q10404.
PT Hetero-polymorphic protein prodn. - by culturing cells transformed
PT with vector contg. structural gene contg. at least one intron
PS Disclosure: Page 12; 19pp; English.
CC The partial FSH protein is encoded by a 2000bp Ddel-Sau3A segment
CC comprising the protein coding region of exon II in addition to 40bp
CC upstream of the start ATG, the protein coding region of exon III
CC and the 1.6kb intron separating the exons. The partial genomic
CC sequence was inserted into the XhoI site of the CLH3AXSV20DC
CC expression vector which contains a murine ornithine decarboxylase
CC coding region. The vector was used in a comparison of
CC engineered cDNA alpha-subunit with the full-length alpha-subunit.
CC See also Q10075 and Q10077.
SQ Sequence 129 AA;

Query Match 53.2%; Score 91; DB 1; Length 129;
Best Local Similarity 44.8%; Pred. No. 1.8e-05;
Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Oy 2 cttyrdfiyrtveipgcpplhvafysypva 30
||::: || || || | :||| |
Db 69 CTFKELVYETVRVPGCAHADSLSLYTPVA 97

RESULT 35
P60618
ID P60618 standard; Protein; 129 AA.
AC P60618;
DE 29-JUL-1991 (first entry)
DE Sequence of mature beta follicle stimulating hormone (FSH)
DE and signal peptide.
KW Ovulation; fertility; diagnosis; pituitary function.
OS Homo sapiens.
FH Key Location/Qualifiers
FT peptide 1..18
FT /label= signal
FT FT 19..129
PN W08604589-A.
PD 14-AUG-1986.
PF 30-JAN-1986; U00223.
PR 02-NOV-1983; US-548228.
PR 30-JAN-1985; US-696647.
PA (REDD/) REDDY V B.
PA (INTE-) INTEGRATED GENETICS.
PI Reddy VB, Hsiung N, Beck AK, Bernstein EG;
DR WPI: 86-225447/34.
DR N-PSDB: N60616.
PT New active hetero-dimeric human follicle stimulating hormone - is
PT obtd. from single culture of transformed cells and
PT post-translational modification may occur
PS Example; Page 10; 25pp; English.
CC The beta-subunit of human FSH synthesised by a cell comprising an
CC expression vector comprising heterologous DNA encoding the beta-
CC subunit of FSH. The cells are mammalian. For prodn. of the beta-
CC subunit, the vector comprises at least the 69% transforming region
CC of the bovine papilloma virus genome. The cell is a mouse cell and
CC the expression vector comprises a plasmid, e.g. deposited as NRRL
CC B-15923.
SQ Sequence 129 AA;

Query Match 53.2%; Score 91; DB 1; Length 129;
Best Local Similarity 44.8%; Pred. No. 1.8e-05;
Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Oy 2 cttyrdfiyrtveipgcpplhvafysypva 30
||::: || || || | :||| |
Db 69 CTFKELVYETVRVPGCAHADSLSLYTPVA 97

```

```

RESULT 36
R30786
ID R30786 standard; Protein; 129 AA.
AC R30786;
DE 26-MAY-1993 (first entry)
DE Follicle stimulating hormone beta.
KW FSH; human; dimer; hormone; reproduction; modification.
OS Homo sapiens.
FH Key Location/Qualifiers
FT peptide 1..18
FT /note= "signal peptide"
FT protein 19..129
FT /note= "mature FSH beta"
FT US5177193-A.
PN 05-JAN-1993.
PD 21-FEB-1989; 313646.
PR 21-FEB-1989; US-313646.
PR 01-JUN-1990; US-532254.
PA (UNIW ) UNIV WASHINGTON.
PI Boime I, Matzuk MM;
DR WPI: 93-035777/04.
DR N-PSDB: Q34971.
PT Modified luteinising hormone beta-sub-unit - shows enhanced
PT ability to dimerise and enhanced secretion as a dimer from
PT recombinant host cells
PS Disclosure; Fig 2; 26pp; English.
CC Genomic DNA from JAR choriocarcinoma cells (a human placental donor)
CC was used to create a library which was screened by a 41 mer corresp.
CC to amino acids 94-107 of exon III of human FSH beta, and positive
CC clones isolated. Sequencing revealed that the coding sequence is
CC divided into three exons. Exon I contains a 5' untranslated tract
CC previously reported to encode two transcripts of either 33 or 63 bp.
CC Exon II encodes an 18 amino acid signal peptide and amino acids 1-35
CC of the mature protein. Exon III encodes amino acids 36-111 and ca.
CC 1.1 kb of 3' untranslated sequence. Exons I and II are separated by
CC an intron of 800bp, and exons II and III by an intron of 1.4 kb. The
CC notation "nnn" in the sequence represents gaps in the sequence
CC given in the specification, of 0.7 kb in the first intron, 1.35 kb in
CC the second intron and of 0.55 kb in the 3' untranslated sequence.
SQ Sequence 129 AA;

```

```

Query Match 53.2%; Score 91; DB 1; Length 129;
Best Local Similarity 44.8%; Pred. No. 1.8e-05;
Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

```

```

Oy 2 cttyrdfiyrtveipgcpplhvafysypva 30
||::: || || || | :||| |
Db 69 CTFKELVYETVRVPGCAHADSLSLYTPVA 97

```

```

RESULT 37
R99418
ID R99418 standard; Protein; 129 AA.
AC R99418;
DE 18-JAN-1997 (first entry)
DE FSH beta subunit.
KW Heteropolymorphic protein; dimeric glycoprotein; hormone;
KW follicle stimulating hormone; FSH.
OS Homo sapiens.
PN EP-735139-A1.
PD 02-OCT-1996.
PF 20-JUN-1990; 306743.
PR 20-JUN-1989; US-368628.
PA (GENZ ) GENZYME CORP.
PI Chappel SC, Kelton CA, Nugent NP;
DR WPI: 96-478414/48.
DR N-PSDB: T41696.
PT Increased recombinant protein production - esp. heteropolymorphic
PT glycoprotein hormones, uses DNA contg. at least one intron
PS Example 1; Page 16-17; 26pp; English.
CC The beta subunit (R99418) of human follicle stimulating hormone

```


KW peptide hormone; chorionic gonadotropin.
 OS Homo sapiens. Location/Qualifiers
 FH Peptide i. 18
 FT /note= "signal peptide"
 FT 19..129
 FT Protein /note= "mature FSH-beta"
 FT US5712122-A.
 PN 27-JAN-1998.
 PD 07-JUN-1995; 483023.
 PF 01-JUN-1990; US-532254.
 PR 21-FEB-1989; US-313646.
 PR 24-SEP-1992; US-950801.
 PR 06-MAY-1994; US-239256.
 PR 07-JUN-1995; US-483023.
 PA (UNIV) UNIV WASHINGTON.
 PI Boime I, Keene JL, Matzuk MM;
 DR WPI; 98-119981/11.
 DR N-PSDB; V19789.
 PT Modified protein extended with carboxy-terminal peptide - retains
 PT activity of parent protein which is human chorionic gonadotropin
 PS Disclosure; Fig 2; 28pp; English.
 CC This sequence is the human follicle-stimulating hormone (FSH) beta
 CC chain. The protein can be modified to give a protein of the invention
 CC that retains the biological activity of the parent protein. The modified
 CC protein comprises an extension of the carboxy terminus of the parent
 CC protein with an amino acid sequence of the carboxy-terminal peptide (CTP)
 CC of human chorionic gonadotropin (HCG) beta-subunit, consisting of amino
 CC acids 112-118 to amino acids 145 of HCG, or a variant that differs from
 CC the CTP by 1-5 conservative amino acid substitutions. The parent protein
 CC is preferably a peptide hormone, especially FSH. A fusion protein
 CC comprising FSH beta chain fused to the CTP of HCG is very efficiently
 CC produced and secreted and may have an extended half-life in vivo.
 SQ Sequence 129 AA;

Query Match 53.2%; Score 91; DB 1; Length 129;
 Best Local Similarity 44.8%; Pred. No. 1.8e-05;
 Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 2 cttydfiyrtveipgpcplhvapyfsypva 30
 ||:::| ||:| || | :|||
 Db 69 CTFKELVYETVRVPGCAHADSLSLYTPVA 97

RESULT 41

W62027
 ID W62027 standard; Protein; 129 AA.
 AC W62027;
 DT 10-SEP-1998 (first entry)
 DE Human beta follicle stimulating hormone.
 KW Human; beta follicle stimulating hormone; FSH; fertility treatment;
 KW heterodimeric; chorionic gonadotropin; LH; luteinising hormone.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc_difference 128
 FT /note= "encoded by AAA"
 FT Misc_difference 129
 FT /note= "encoded by GAA"
 FT US5767251-A.
 PN 16-JUN-1998.
 PD 22-JAN-1993; 008233.
 PF 27-APR-1990; US-515481.
 PR 02-NOV-1983; US-548228.
 PR 30-JAN-1985; US-696647.
 PR 15-MAR-1989; US-323772.
 PR 22-JAN-1993; US-008233.
 PR (GENZ) GENZYME CORP.
 PA Beck AK, Bernstein EG, Hsiung N, Reddy VB;
 DR WPI; 98-361787/31.
 DR N-PSDB; V37929.
 PT Recombinantly produced human fertility hormones - free from human
 PT protein contaminants, used for fertility treatment

PS Disclosure; Column 13-14; 23pp; English.
 CC The present invention describes new recombinantly produced
 CC biologically active heterodimeric active human fertility hormones
 CC selected from one of: human chorionic gonadotropin (hCG), human
 CC luteinising hormone (hLH) and human follicle stimulating hormone
 CC (hFSH) which are free from contamination by any other proteins of
 CC human origin. The present sequence represents human beta-FSH from the
 CC present invention. The hormones are used in human fertility treatment
 CC and the FSH can be especially used in diagnostic tests for fertility
 CC and pituitary functions. The hormones are free of contaminating
 CC proteins that would normally be found in preparations from natural
 CC sources.
 SQ Sequence 129 AA;

Query Match 53.2%; Score 91; DB 1; Length 129;
 Best Local Similarity 44.8%; Pred. No. 1.8e-05;
 Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 2 cttydfiyrtveipgpcplhvapyfsypva 30
 ||:::| ||:| || | :|||
 Db 69 CTFKELVYETVRVPGCAHADSLSLYTPVA 97

RESULT 42

W89411
 ID W89411 standard; Protein; 129 AA.
 AC W89411;
 DT 10-MAR-1999 (first entry)
 DE Human follicle stimulating hormone beta-subunit.
 KW Human; follicle stimulating hormone beta-subunit; beta FSH;
 KW luteinising hormone; LH; fertility.
 OS Homo sapiens.
 PN US5856137-A.
 PD 05-JAN-1999.
 PF 07-JUN-1995; 485503.
 PR 27-APR-1990; US-515481.
 PR 02-NOV-1983; US-548228.
 PR 30-JAN-1985; US-696647.
 PR 15-MAR-1989; US-323772.
 PR 30-JUN-1994; US-268734.
 PR 07-JUN-1995; US-485503.
 PA (GENZ) GENZYME CORP.
 PI Beck AK, Bernstein EG, Hsiung N, Reddy VB;
 DR WPI; 99-105109/09.
 DR N-PSDB; V82015.
 PT Isolated DNA encoding luteinising hormone beta subunit - useful with
 PT DNA encoding the alpha subunit for the recombinant expression of
 PT biologically active human luteinising hormone
 PS Disclosure; Column 13-14; 26pp; English.
 CC The present invention describes isolated DNA encoding the beta-subunit
 CC of human luteinising hormone (hLH), and having the sequence encoding the
 CC hLH which is present in Escherichia coli clone pCL28XhLHbVP (ATCC
 CC 39475). Products from the present invention can be used to produce the
 CC beta-subunit of hLH. A vector containing the DNA encoding the beta-
 CC subunit of hLH can be used with an alpha-subunit encoding sequence to
 CC produce glycosylated, biologically active, heterodimeric hLH for use in
 CC conditions associated with human fertility. The present sequence
 CC represents the human beta-subunit of follicle stimulating hormone
 CC (FSH), from the present invention.
 SQ Sequence 129 AA;

Query Match 53.2%; Score 91; DB 1; Length 129;
 Best Local Similarity 44.8%; Pred. No. 1.8e-05;
 Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 2 cttydfiyrtveipgpcplhvapyfsypva 30
 ||:::| ||:| || | :|||
 Db 69 CTFKELVYETVRVPGCAHADSLSLYTPVA 97

RESULT 43

W99543
ID W99543 standard; Protein; 129 AA.
AC W99543:
DT 08-JUN-1999 (first entry)
DE hFSH-beta analogue hFSH-beta-v31C.
KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
KW human follicle stimulating hormone; human thyroid stimulating hormone;
KW stability; primer; amplification; PCR; mutation.
OS Homo sapiens.
OS Synthetic.
PN W09858957-A2.
PD 30-DEC-1998.
PF 25-JUN-1998; U13070.
PR 25-JUN-1997; US-050784.
PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PA (MCIN-) MCINNIS P G.
PI Moyle WR;
DR WPI; 99-081219/07.
PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
PT or hTSH, have an intersubunit disulphide crosslink between the
PT alpha- and beta-subunits to improve stability
PS Disclosure; Fig 32B; 139pp; English.
CC The invention relates to the production of analogues of a heterodimeric
CC subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
CC (hCG), human luteinising hormone (hLH), human follicle stimulating
CC hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
CC muteins, which are modified to contain an intersubunit disulphide bond,
CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
CC improved stability, the analogue retaining at least a portion of the
CC bioactivity for the corresponding native GPH receptor. This sequence
CC represents a mutant hFSH-beta subunit used for the generation of
CC the modified GPHs. The improved analogues are designed specifically
CC to reduce perturbation of the 3-dimensional structure of the hormone,
CC thereby creating greater likelihood that the dimer will be formed in vivo
CC and the formed dimer will have affinity for the native receptors and have
CC agonistic activity on them. The changes stabilise the GPHs and prolong
CC as for the native GPHs.
SQ Sequence 129 AA;

Query Match 53.2%; Score 91; DB 1; Length 129;
Best Local Similarity 44.8%; Pred. No. 1.8e-05;
Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 2 ctgydfiyrvtveipgcpvhvafsyypva 30
||::: || || || | :|||
Db 69 CTFKELVYETVRVPGCAHHAHDSLYTPVA 97

RESULT 45
W99505
ID W99505 standard; Protein; 129 AA.
AC W99505;
DT 08-JUN-1999 (first entry)
DE Glycoprotein hormone analogue hFSH-beta-BQ48C.
KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
KW human follicle stimulating hormone; human thyroid stimulating hormone;
KW stability; primer; amplification; PCR; mutation.
OS Homo sapiens.
OS Synthetic.
PN W09858957-A2.
PD 30-DEC-1998.
PF 25-JUN-1998; U13070.
PR 25-JUN-1997; US-050784.
PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PA (MCIN-) MCINNIS P G.
PI Moyle WR;
DR WPI; 99-081219/07.
PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
PT or hTSH, have an intersubunit disulphide crosslink between the
PT alpha- and beta-subunits to improve stability
PS Example 12; Page 88; 139pp; English.
CC The invention relates to the production of analogues of a heterodimeric
CC subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
CC (hCG), human luteinising hormone (hLH), human follicle stimulating
CC hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
CC muteins, which are modified to contain an intersubunit disulphide bond,
CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
CC improved stability, the analogue retaining at least a portion of the
CC bioactivity for the corresponding native GPH receptor. This sequence
CC represents a mutant hFSH-beta subunit used for the generation of the
CC modified GPHs. The improved analogues are designed specifically to
CC reduce perturbation of the 3-dimensional structure of the hormone,
CC thereby creating greater likelihood that the dimer will be formed in vivo
CC and the formed dimer will have affinity for the native receptors and have
CC agonistic activity on them. The changes stabilise the GPHs and prolong
CC as for the native GPHs.
SQ Sequence 129 AA;

Query Match 53.2%; Score 91; DB 1; Length 129;
Best Local Similarity 44.8%; Pred. No. 1.8e-05;
Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 2 ctgydfiyrvtveipgcpvhvafsyypva 30
||::: || || || | :|||
Db 69 CTFKELVYETVRVPGCAHHAHDSLYTPVA 97

RESULT 44
W99546
ID W99546 standard; Protein; 129 AA.
AC W99546;
DT 08-JUN-1999 (first entry)
DE hFSH-beta analogue hFSH-beta-52C.
KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
KW human follicle stimulating hormone; human thyroid stimulating hormone;
KW stability; primer; amplification; PCR; mutation.
OS Homo sapiens.
OS Synthetic.
PN W09858957-A2.
PD 30-DEC-1998.
PF 25-JUN-1998; U13070.
PR 25-JUN-1997; US-050784.
PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PA (MCIN-) MCINNIS P G.
PI Moyle WR;
DR WPI; 99-081219/07.
PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
PT or hTSH, have an intersubunit disulphide crosslink between the


```
Query Match      53.2%; Score 91; DB 1; Length 129;
Best Local Similarity 44.8%; Pred. No. 1.8e-05;
Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 2 cttydfiyrvtveipgcplhvafysypva 30
   ||::: || || || || | :|||
Db 69 CTFKELVYETVRVPGCAHADSLSYTPVA 97

RESULT 46
ID W99499 standard; Protein: 129 AA.
AC W99499;
DT 08-JUN-1999 (first entry)
DE Glycoprotein hormone analogue hFSH-beta-T92C.
KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
KW human follicle stimulating hormone; human thyroid stimulating hormone;
KW stability; primer; amplification; PCR; mutation.
OS Homo sapiens.
OS Synthetic.
PN W09858957-A2.
PD 30-DEC-1998.
PF 25-JUN-1998; U13070.
PR (ISTF ) ARS APPLIED RES SYSTEMS HOLDING NV.
PA (MCIN-) MCINNIS P G.
PI Moyle WR;
DR WPI: 99-081219/07.
PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
PT or hTSH, have an intersubunit disulphide crosslink between the
PT alpha- and beta-subunits to improve stability
PS Example 12; Page 88; 139pp; English.
CC The invention relates to the production of analogues of a heterodimeric
CC hormone (hFSH), human thyroid stimulating hormone (hLH), human chorionic gonadotropin
CC subunit glycoprotein hormone (GPH) e.g. human follicle stimulating
CC (hCG), human luteinising hormone (hLH), human follicle stimulating
CC hormone (hFSH), heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
CC muteins, which are modified to contain an intersubunit disulphide bond,
CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
CC improved stability, the analogue retaining at least a portion of the
CC bioactivity for the corresponding native GPH receptor. This sequence
CC represents a mutated hFSH-beta subunit used for the generation of the
CC modified GPHs. The improved analogues are designed specifically to
CC reduce perturbation of the 3-dimensional structure of the hormone,
CC thereby creating greater likelihood that the dimer will be formed in vivo
CC and the formed dimer will have affinity for the native receptors and have
CC agonistic activity on them. The changes stabilise the GPHs and prolong
CC the biological activities of the hormones. The analogues can have uses
CC as for the native GPHs.
SQ Sequence 129 AA;

Query Match      53.2%; Score 91; DB 1; Length 129;
Best Local Similarity 44.8%; Pred. No. 1.8e-05;
Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 2 cttydfiyrvtveipgcplhvafysypva 30
   ||::: || || || || | :|||
Db 69 CTFKELVYETVRVPGCAHADSLSYTPVA 97

RESULT 48
ID W99501 standard; Protein: 129 AA.
AC W99501;
DT 08-JUN-1999 (first entry)
DE Glycoprotein hormone analogue hFSH-beta-T34C.
KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
KW human follicle stimulating hormone; human thyroid stimulating hormone;
KW stability; primer; amplification; PCR; mutation.
OS Homo sapiens.
OS Synthetic.
PN W09858957-A2.
PD 30-DEC-1998.
PF 25-JUN-1998; U13070.
PR (ISTF ) ARS APPLIED RES SYSTEMS HOLDING NV.
PA (MCIN-) MCINNIS P G.
PI Moyle WR;
DR WPI: 99-081219/07.
PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
PT or hTSH, have an intersubunit disulphide crosslink between the
PT alpha- and beta-subunits to improve stability
PS Example 12; Page 88; 139pp; English.
CC The invention relates to the production of analogues of a heterodimeric
CC subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
CC (hCG), human luteinising hormone (hLH), human follicle stimulating
CC hormone (hFSH), heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
CC muteins, which are modified to contain an intersubunit disulphide bond,
CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
CC improved stability, the analogue retaining at least a portion of the
CC bioactivity for the corresponding native GPH receptor. This sequence
CC represents a mutated hFSH-beta subunit used for the generation of the
CC modified GPHs. The improved analogues are designed specifically to
CC reduce perturbation of the 3-dimensional structure of the hormone,
CC thereby creating greater likelihood that the dimer will be formed in vivo
CC and the formed dimer will have affinity for the native receptors and have
CC agonistic activity on them. The changes stabilise the GPHs and prolong
CC the biological activities of the hormones. The analogues can have uses
CC as for the native GPHs.
SQ Sequence 129 AA;
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 14, 2000, 08:55:46 ; Search time 35.12 Seconds
(without alignments)
12.333 Million cell updates/sec

Title: BETA-CHAIN
Perfect score: 171
Sequence: 1 vctyrdfiyrtveipgclhvpvysypva 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 145308 seqs, 14437401 residues

Total number of hits satisfying chosen parameters: 145308

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 99%

Listing first 1000 summaries

Database : Issued Patents.AA.*
1: /cgnl_6/ptodata/1/iaa/5A-COMB.pep.*
2: /cgnl_6/ptodata/1/iaa/5B-COMB.pep.*
3: /cgnl_6/ptodata/1/iaa/5-COMB.pep.*
4: /cgnl_6/ptodata/1/iaa/PCITUS-COMB.pep.*
5: /cgnl_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	163	95.3	112	1	US-08-425-673-6
2	91	53.2	75	1	US-08-239-256-3
3	91	53.2	76	5	5177193-5
4	91	53.2	111	1	US-08-425-673-3
5	91	53.2	111	1	US-08-425-673-4
6	91	53.2	131	5	5177193-3
7	90	52.6	149	1	US-08-425-673-5
8	86	50.3	88	2	US-08-327-362-2
9	86	50.3	114	1	US-08-425-673-7
10	86	50.3	114	1	US-08-425-673-8
11	86	50.3	114	1	US-08-425-673-9
12	86	50.3	117	1	US-08-425-673-11
13	86	50.3	117	1	US-08-425-673-12
14	86	50.3	145	1	US-08-425-673-1
15	86	50.3	145	1	US-08-425-673-2
16	86	50.3	145	1	US-08-425-673-10
17	86	50.3	145	1	US-08-298-189B-1
18	86	50.3	145	1	US-08-475-213-10
19	86	50.3	145	2	US-08-395-238-2
20	83	48.5	98	2	US-08-709-924-23
21	83	48.5	98	2	US-08-709-925-23
22	83	48.5	165	2	US-08-709-924-2
23	83	48.5	165	2	US-08-709-925-2
24	70	40.9	88	2	US-08-709-924-24
25	70	40.9	88	2	US-08-709-925-24
26	62	36.3	26	1	US-08-036-555B-20
27	62	36.3	26	1	US-08-469-569-20
28	62	36.3	26	1	US-08-249-322A-20

29	62	36.3	26	1	US-08-469-526A-20	Sequence 20, Appl
30	62	36.3	26	2	US-08-734-591A-20	Sequence 20, Appl
31	62	36.3	26	2	US-08-469-660-20	Sequence 20, Appl
32	62	36.3	26	4	PCT-US94-05083C-20	Sequence 20, Appl
33	62	36.3	26	4	PCT-US95-06846A-20	Sequence 27, Appl
34	47	27.5	453	2	US-08-599-171A-27	Sequence 27, Appl
35	47	27.5	453	2	US-08-646-590B-27	Sequence 27, Appl
36	47	27.5	453	3	US-09-069-226-27	Sequence 27, Appl
37	46.5	27.2	289	1	US-07-593-657-7	Sequence 7, Appl
38	46	26.9	105	1	US-08-482-577B-27	Sequence 27, Appl
39	46	26.9	105	3	US-08-478-097A-22	Sequence 22, Appl
40	46	26.9	121	1	US-08-481-377-20	Sequence 18, Appl
41	46	26.9	121	2	US-08-491-835-18	Sequence 18, Appl
42	46	26.9	121	3	US-05-153-733A-20	Sequence 20, Appl
43	46	26.9	121	4	PCT-US94-00666-20	Sequence 20, Appl
44	46	26.9	121	4	PCT-US94-00685-18	Sequence 18, Appl
45	46	26.9	122	1	US-08-581-529B-16	Sequence 16, Appl
46	46	26.9	122	1	US-08-455-559-22	Sequence 22, Appl
47	46	26.9	122	2	US-08-525-596B-26	Sequence 26, Appl
48	46	26.9	122	2	US-08-581-528A-16	Sequence 16, Appl
49	46	26.9	122	4	PCT-US94-00657-22	Sequence 22, Appl
50	46	26.9	122	4	PCT-US94-07762-16	Sequence 16, Appl
51	46	26.9	122	4	PCT-US94-07799-16	Sequence 16, Appl
52	46	26.9	351	1	US-08-197-792-39	Sequence 39, Appl
53	46	26.9	351	1	US-08-459-850-39	Sequence 39, Appl
54	46	26.9	351	1	US-08-459-214-39	Sequence 39, Appl
55	45.5	26.6	533	1	US-08-220-151-15	Sequence 15, Appl
56	45.5	26.6	533	1	US-08-413-118-15	Sequence 15, Appl
57	45.5	26.6	533	3	US-08-473-446-15	Sequence 15, Appl
58	45	26.3	348	2	US-08-185-828A-14	Sequence 14, Appl
59	44.5	26.0	84	2	US-08-744-670-5	Sequence 5, Appl
60	44.5	26.0	84	2	US-08-149-933-5	Sequence 5, Appl
61	43	25.1	471	2	US-08-399-889-24	Sequence 24, Appl
62	43	25.1	471	3	US-09-167-364-24	Sequence 24, Appl
63	43	25.1	1128	1	US-08-111-939-2	Sequence 2, Appl
64	43	25.1	2254	2	US-08-677-010-3	Sequence 3, Appl
65	43	25.1	2254	2	US-08-790-519-3	Sequence 3, Appl
66	42.5	24.9	351	2	US-08-828-488-5	Sequence 5, Appl
67	42.5	24.9	398	2	US-08-599-171A-29	Sequence 29, Appl
68	42.5	24.9	398	2	US-08-646-590B-29	Sequence 29, Appl
69	42.5	24.9	398	3	US-09-069-226-29	Sequence 29, Appl
70	42.5	24.9	476	2	US-08-828-488-3	Sequence 3, Appl
71	42.5	24.9	477	2	US-08-828-488-1	Sequence 1, Appl
72	42	24.6	680	2	US-08-674-351-2	Sequence 2, Appl
73	42	24.6	1658	2	US-08-609-049A-13	Sequence 13, Appl
74	42	24.6	1726	2	US-08-609-049A-30	Sequence 30, Appl
75	41	24.0	92	2	US-08-395-238-1	Sequence 1, Appl
76	41	24.0	116	2	US-08-599-895-5	Sequence 5, Appl
77	41	24.0	191	2	US-08-230-665A-161	Sequence 161, App
78	41	24.0	191	4	PCT-US93-10398-161	Sequence 12, Appl
79	41	24.0	334	1	US-08-182-025-12	Sequence 12, Appl
80	41	24.0	334	2	US-08-185-828A-16	Sequence 16, Appl
81	41	24.0	365	2	US-08-846-762-80	Sequence 80, Appl
82	41	24.0	365	2	US-08-576-626A-46	Sequence 46, Appl
83	41	24.0	489	3	US-08-369-822C-27	Sequence 27, Appl
84	41	24.0	656	1	US-08-444-005-15	Sequence 15, Appl
85	40.5	23.7	266	1	US-08-202-857-2	Sequence 2, Appl
86	40.5	23.7	873	3	US-09-187-331-6	Sequence 6, Appl
87	40.5	23.7	925	2	US-08-392-946-1	Sequence 1, Appl
88	40.5	23.7	925	2	US-08-504-169-1	Sequence 1, Appl
89	40.5	23.7	925	2	PCT-US94-14893-1	Sequence 25, Appl
90	40	23.4	218	2	US-08-399-889-25	Sequence 25, Appl
91	40	23.4	218	3	US-09-167-364-25	Sequence 25, Appl
92	40	23.4	371	2	US-08-837-593-8	Sequence 8, Appl
93	40	23.4	550	1	US-08-484-453-2	Sequence 2, Appl
94	40	23.4	550	1	US-08-484-494-2	Sequence 2, Appl
95	40	23.4	550	2	US-08-454-212-2	Sequence 2, Appl
96	40	23.4	707	1	US-08-528-122-18	Sequence 18, Appl
97	40	23.4	707	4	PCT-US95-11720-18	Sequence 18, Appl
98	40	23.4	750	1	US-08-325-553-2	Sequence 2, Appl
99	40	23.4	750	2	US-08-394-152A-2	Sequence 2, Appl
100	40	23.4	758	1	US-08-258-188-2	Sequence 2, Appl
101	40	23.4	758	1	US-08-526-813-2	Sequence 2, Appl

102	40	23.4	758	4	PCT-US95-08554-2	Sequence 2, Appli	175	37	21.6	261	3	US-08-483-857-10	Sequence 10, Appli
103	40	23.4	875	1	US-08-785-241-5	Sequence 5, Appli	176	37	21.6	276	1	US-07-952-817-27	Sequence 27, Appli
104	40	23.4	1266	1	US-08-468-557-4	Sequence 4, Appli	177	37	21.6	288	2	US-08-737-825-8	Sequence 8, Appli
105	40	23.4	1548	1	US-08-463-092B-7	Sequence 7, Appli	178	37	21.6	377	2	US-08-823-986A-4	Sequence 4, Appli
106	40	23.4	1548	2	US-08-460-907B-7	Sequence 7, Appli	179	37	21.6	377	3	US-08-784-385-4	Sequence 4, Appli
107	39.5	23.1	125	2	US-08-888-497-42	Sequence 42, Appli	180	37	21.6	377	4	PCT-US95-10403-4	Sequence 4, Appli
108	39.5	23.1	125	4	PCT-US94-07926-42	Sequence 42, Appli	181	37	21.6	403	5	US-08-955-10403-4	Sequence 4, Appli
109	39.5	23.1	128	3	US-08-804-439A-2	Sequence 2, Appli	182	37	21.6	473	1	US-08-073-383-6	Sequence 6, Appli
110	39.5	23.1	128	3	US-08-720-229-2	Sequence 2, Appli	183	37	21.6	473	1	PCT-US94-06365-6	Sequence 6, Appli
111	39.5	23.1	146	2	US-08-888-497-35	Sequence 35, Appli	184	37	21.6	567	1	US-08-261-822A-12	Sequence 12, Appli
112	39.5	23.1	146	4	PCT-US94-07926-35	Sequence 35, Appli	185	37	21.6	567	1	PCT-US95-07744A-12	Sequence 12, Appli
113	39.5	23.1	438	2	US-08-282-197C-59	Sequence 59, Appli	186	37	21.6	582	1	US-08-261-086-2	Sequence 2, Appli
114	39.5	23.1	486	2	US-08-821-355A-8	Sequence 8, Appli	187	37	21.6	582	1	US-08-261-086-4	Sequence 4, Appli
115	39.5	23.1	486	2	US-09-003-687A-8	Sequence 8, Appli	188	37	21.6	582	1	US-08-261-086-6	Sequence 6, Appli
116	39.5	23.1	511	2	US-08-821-355A-9	Sequence 9, Appli	189	37	21.6	811	2	US-08-933-750C-45	Sequence 45, Appli
117	39.5	23.1	511	2	US-09-003-687A-9	Sequence 9, Appli	190	37	21.6	816	2	US-08-267-803B-9	Sequence 9, Appli
118	39.5	23.1	513	1	US-08-464-266-2	Sequence 2, Appli	191	37	21.6	1073	1	US-07-695-564-1	Sequence 1, Appli
119	39.5	23.1	513	1	US-08-464-272-2	Sequence 2, Appli	192	37	21.6	1073	1	US-08-241-387-1	Sequence 1, Appli
120	39.5	23.1	617	1	US-08-279-700-14	Sequence 14, Appli	193	37	21.6	1091	1	US-07-695-564-3	Sequence 3, Appli
121	39.5	23.1	2465	2	US-08-596-291-3	Sequence 3, Appli	194	37	21.6	1091	1	US-08-241-387-3	Sequence 3, Appli
122	39.5	23.1	2466	4	PCT-US94-09943-2	Sequence 2, Appli	195	37	21.6	1385	1	US-07-876-280-2	Sequence 2, Appli
123	39	22.8	309	1	US-07-966-187-1	Sequence 1, Appli	196	37	21.6	1385	1	US-07-675-772-2	Sequence 2, Appli
124	39	22.8	337	1	US-07-966-187-25	Sequence 25, Appli	197	37	21.6	1385	1	US-08-063-170-2	Sequence 2, Appli
125	39	22.8	338	1	US-07-966-187-3	Sequence 3, Appli	198	37	21.6	1385	1	US-08-158-232-2	Sequence 2, Appli
126	39	22.8	338	1	US-07-966-187-16	Sequence 16, Appli	199	37	21.6	1385	1	US-08-304-626-2	Sequence 2, Appli
127	39	22.8	370	1	US-07-966-187-12	Sequence 12, Appli	200	37	21.6	1385	1	US-08-316-301A-2	Sequence 2, Appli
128	39	22.8	370	1	US-07-966-187-14	Sequence 14, Appli	201	37	21.6	1385	2	US-08-611-928-2	Sequence 2, Appli
129	39	22.8	370	1	US-07-966-187-18	Sequence 18, Appli	202	37	21.6	1385	2	PCT-US92-03624-2	Sequence 2, Appli
130	39	22.8	647	1	US-08-218-943-1	Sequence 1, Appli	203	37	21.6	1385	5	US-08-460-751-2	Sequence 2, Appli
131	39	22.8	933	2	US-08-313-200-1	Sequence 1, Appli	204	37	21.6	1385	5	US-08-460-751-2	Sequence 2, Appli
132	39	22.8	933	4	PCT-US93-03837-1	Sequence 1, Appli	205	37	21.6	1385	5	US-08-460-751-2	Sequence 2, Appli
133	39	22.8	2647	2	US-08-593-562B-8	Sequence 8, Appli	206	36.5	21.3	128	3	US-08-804-439A-4	Sequence 4, Appli
134	39	22.8	2647	2	US-08-779-113-8	Sequence 8, Appli	207	36.5	21.3	128	3	US-08-720-229-4	Sequence 4, Appli
135	38.5	22.5	178	2	US-08-680-326-4	Sequence 4, Appli	208	36.5	21.3	175	2	US-08-809-267-5	Sequence 5, Appli
136	38.5	22.5	192	1	US-08-086-428B-92	Sequence 92, Appli	209	36.5	21.3	175	4	PCT-US95-13662A-5	Sequence 5, Appli
137	38.5	22.5	192	2	US-08-468-570-92	Sequence 92, Appli	210	36.5	21.3	237	1	US-08-910-973-13	Sequence 13, Appli
138	38.5	22.5	192	2	US-08-290-665A-92	Sequence 92, Appli	211	36.5	21.3	263	2	US-08-353-476-66	Sequence 66, Appli
139	38.5	22.5	192	4	PCT-US95-10398-92	Sequence 92, Appli	212	36.5	21.3	289	2	US-08-580-545B-4	Sequence 4, Appli
140	38.5	22.5	443	2	US-08-833-963C-2	Sequence 2, Appli	213	36.5	21.3	306	3	US-08-301-162-4	Sequence 4, Appli
141	38.5	22.5	537	3	US-08-920-610-2	Sequence 2, Appli	214	36.5	21.3	320	1	US-08-726-525-4	Sequence 4, Appli
142	38.5	22.5	837	2	US-08-583-326-117	Sequence 117, Appli	215	36.5	21.3	320	2	US-08-487-942-4	Sequence 4, Appli
143	38.5	22.5	858	2	US-08-583-562B-2	Sequence 2, Appli	216	36.5	21.3	320	2	US-08-726-036A-4	Sequence 4, Appli
144	38.5	22.5	1711	3	US-08-369-822C-10	Sequence 10, Appli	217	36.5	21.3	331	1	US-08-356-180-3	Sequence 3, Appli
145	38.5	22.5	4544	1	US-08-469-486-52	Sequence 52, Appli	218	36.5	21.3	334	1	US-08-118-270-22	Sequence 22, Appli
146	38.5	22.5	4544	2	US-08-469-658-52	Sequence 52, Appli	219	36.5	21.3	334	4	PCT-US93-08528-22	Sequence 22, Appli
147	38	22.2	46	1	US-07-946-054-12	Sequence 12, Appli	220	36.5	21.3	380	1	US-08-585-758A-4	Sequence 4, Appli
148	38	22.2	46	4	PCT-US93-08638-12	Sequence 12, Appli	221	36.5	21.3	380	1	US-08-977-818-4	Sequence 4, Appli
149	38	22.2	293	2	US-08-489-141A-2	Sequence 2, Appli	222	36.5	21.3	380	2	US-08-670-274B-4	Sequence 4, Appli
150	38	22.2	293	2	US-08-576-626A-59	Sequence 59, Appli	223	36.5	21.3	390	2	US-08-786-999-1	Sequence 1, Appli
151	38	22.2	364	2	US-08-483-926A-10	Sequence 10, Appli	224	36.5	21.3	397	2	US-08-750-134A-9	Sequence 9, Appli
152	38	22.2	364	2	US-08-737-045-10	Sequence 10, Appli	225	36.5	21.3	452	3	US-08-301-162-20	Sequence 20, Appli
153	38	22.2	769	2	US-08-789-078-1	Sequence 1, Appli	226	36.5	21.3	460	1	US-08-351-981-4	Sequence 4, Appli
154	38	22.2	769	2	US-08-752-633-1	Sequence 1, Appli	227	36.5	21.3	460	1	US-08-351-981-9	Sequence 9, Appli
155	38	22.2	769	2	US-08-476-062A-45	Sequence 45, Appli	228	36.5	21.3	505	1	US-08-660-765A-2	Sequence 2, Appli
156	38	22.2	769	2	US-07-728-215-31	Sequence 31, Appli	229	36.5	21.3	582	2	US-08-911-445-3	Sequence 3, Appli
157	38	22.2	769	4	PCT-US95-04886-1	Sequence 1, Appli	230	36.5	21.3	845	3	US-08-804-439A-94	Sequence 94, Appli
158	38	22.2	769	4	PCT-US96-01314-45	Sequence 45, Appli	231	36.5	21.3	845	3	US-08-720-229-94	Sequence 94, Appli
159	38	22.2	843	2	US-09-172-977-3	Sequence 3, Appli	232	36.5	21.3	887	1	US-07-596-467-2	Sequence 2, Appli
160	38	22.2	1720	2	US-08-477-451-12	Sequence 12, Appli	233	36.5	21.3	887	1	US-07-934-374-2	Sequence 2, Appli
161	38	22.2	1839	2	US-09-172-977-4	Sequence 4, Appli	234	36.5	21.3	887	1	US-07-783-861C-4	Sequence 4, Appli
162	37.5	21.9	252	2	US-07-857-224B-26	Sequence 26, Appli	235	36.5	21.3	1168	1	US-08-448-170-6	Sequence 6, Appli
163	37.5	21.9	387	2	US-08-872-302-5	Sequence 5, Appli	236	36	21.1	11	4	PCT-US94-07644A-43	Sequence 43, Appli
164	37	21.6	32	5	US-08-428-415-12	Sequence 12, Appli	237	36	21.1	81	2	US-08-744-670-4	Sequence 4, Appli
165	37	21.6	75	1	US-08-379-685-12	Sequence 12, Appli	238	36	21.1	81	2	US-09-149-933-4	Sequence 4, Appli
166	37	21.6	75	1	US-08-854-029-12	Sequence 12, Appli	239	36	21.1	87	1	US-07-918-953-17	Sequence 17, Appli
167	37	21.6	75	2	US-08-253-155A-35	Sequence 35, Appli	240	36	21.1	87	1	US-08-081-661-17	Sequence 17, Appli
168	37	21.6	127	1	US-08-722-126A-5	Sequence 5, Appli	241	36	21.1	150	2	US-08-867-676-3	Sequence 3, Appli
169	37	21.6	188	3	PCT-US95-04258-5	Sequence 5, Appli	242	36	21.1	159	3	US-08-481-435-13	Sequence 13, Appli
170	37	21.6	188	4	US-08-428-415-7	Sequence 7, Appli	243	36	21.1	163	2	US-08-867-676-1	Sequence 1, Appli
171	37	21.6	208	1	US-08-379-685-7	Sequence 7, Appli	244	36	21.1	192	1	US-08-086-428B-62	Sequence 62, Appli
172	37	21.6	208	1	US-08-854-029-7	Sequence 7, Appli	245	36	21.1	192	1	US-08-086-428B-71	Sequence 71, Appli
173	37	21.6	208	2	US-08-436-748-10	Sequence 10, Appli	246	36	21.1	192	1	US-08-086-428B-99	Sequence 99, Appli
174	37	21.6	260	2			247	36	21.1	192	2	US-08-468-570-62	Sequence 62, Appli

248	36	21.1	192	2	US-08-468-570-71	Sequence 71, Appl	321	35.5	20.8	191	4	PCT-US95-10398-181	Sequence 181, App
249	36	21.1	192	2	US-08-468-570-99	Sequence 99, Appl	322	35.5	20.8	191	4	PCT-US95-10398-184	Sequence 184, App
250	36	21.1	192	2	US-08-290-665A-62	Sequence 62, Appl	323	35.5	20.8	292	2	US-08-973-461A-4	Sequence 4, Appl
251	36	21.1	192	2	US-08-290-665A-71	Sequence 71, Appl	324	35.5	20.8	292	2	US-08-648-010-4	Sequence 4, Appl
252	36	21.1	192	2	US-08-290-665A-99	Sequence 99, Appl	325	35.5	20.8	292	2	US-08-852-481-2	Sequence 2, Appl
253	36	21.1	192	2	PCT-US95-10398-62	Sequence 62, Appl	326	35.5	20.8	417	1	US-08-351-981-6	Sequence 6, Appl
254	36	21.1	192	4	PCT-US95-10398-71	Sequence 71, Appl	327	35.5	20.8	417	1	US-08-351-981-7	Sequence 7, Appl
255	36	21.1	192	4	PCT-US95-10398-99	Sequence 99, Appl	328	35.5	20.8	566	1	US-08-810-116-11	Sequence 11, Appl
256	36	21.1	193	2	US-08-336-031-6	Sequence 6, Appl	329	35.5	20.8	566	2	US-07-930-548A-11	Sequence 11, Appl
257	36	21.1	193	4	PCT-US95-06725-6	Sequence 6, Appl	330	35.5	20.8	849	3	US-08-804-439A-17	Sequence 17, Appl
258	36	21.1	201	3	US-08-369-822C-4	Sequence 4, Appl	331	35.5	20.8	849	3	US-08-720-229-17	Sequence 17, Appl
259	36	21.1	233	1	US-09-151-611-1	Sequence 1, Appl	332	35.5	20.8	905	1	US-08-072-574-2	Sequence 2, Appl
260	36	21.1	261	2	US-07-857-224B-60	Sequence 60, Appl	333	35.5	20.8	1135	2	US-08-469-537A-97	Sequence 97, Appl
261	36	21.1	284	2	US-08-766-439-32	Sequence 32, Appl	334	35.5	20.8	1242	2	US-08-680-326-33	Sequence 33, Appl
262	36	21.1	284	2	US-08-766-439-41	Sequence 41, Appl	335	35	20.5	38	1	US-08-176-500-4	Sequence 4, Appl
263	36	21.1	284	2	US-08-766-439-42	Sequence 42, Appl	336	35	20.5	38	1	US-08-471-052A-4	Sequence 4, Appl
264	36	21.1	284	2	US-08-766-439-43	Sequence 43, Appl	337	35	20.5	38	1	US-08-189-331-4	Sequence 4, Appl
265	36	21.1	284	2	US-08-766-439-44	Sequence 44, Appl	338	35	20.5	38	2	US-08-471-939-4	Sequence 4, Appl
266	36	21.1	284	2	US-08-766-439-45	Sequence 45, Appl	339	35	20.5	38	2	US-08-471-800-4	Sequence 4, Appl
267	36	21.1	328	3	US-08-459-046-2	Sequence 2, Appl	340	35	20.5	38	2	US-08-471-068-4	Sequence 4, Appl
268	36	21.1	352	1	US-08-482-577B-4	Sequence 4, Appl	341	35	20.5	50	3	US-08-180-605-2	Sequence 2, Appl
269	36	21.1	366	2	US-08-690-734A-96	Sequence 96, Appl	342	35	20.5	53	1	US-08-602-010A-30	Sequence 30, Appl
270	36	21.1	366	3	US-08-742-185-98	Sequence 98, Appl	343	35	20.5	53	1	US-08-680-726A-30	Sequence 30, Appl
271	36	21.1	385	1	US-08-036-210-5	Sequence 5, Appl	344	35	20.5	98	2	US-08-341-843B-34	Sequence 34, Appl
272	36	21.1	385	2	US-08-449-609-5	Sequence 5, Appl	345	35	20.5	98	2	US-08-427-497E-39	Sequence 39, Appl
273	36	21.1	397	2	US-08-371-377-19	Sequence 19, Appl	346	35	20.5	100	1	US-08-189-199A-3	Sequence 3, Appl
274	36	21.1	398	2	US-08-371-377-17	Sequence 17, Appl	347	35	20.5	101	1	US-08-189-199A-4	Sequence 4, Appl
275	36	21.1	411	1	US-08-381-433A-4	Sequence 4, Appl	348	35	20.5	106	1	US-08-482-577B-24	Sequence 24, Appl
276	36	21.1	411	1	US-08-381-433A-8	Sequence 8, Appl	349	35	20.5	110	2	US-08-828-009-2	Sequence 2, Appl
277	36	21.1	411	2	US-08-336-031-2	Sequence 2, Appl	350	35	20.5	116	5	5206160-1	Patent No. 5206160
278	36	21.1	411	4	US-08-902-853-7	Sequence 7, Appl	351	35	20.5	118	2	US-08-888-497-40	Sequence 40, Appl
279	36	21.1	411	4	PCT-US95-06725-2	Sequence 2, Appl	352	35	20.5	118	4	PCT-US94-07926-40	Sequence 40, Appl
280	36	21.1	446	1	US-07-952-800-4	Sequence 4, Appl	353	35	20.5	120	5	5244657-8	Patent No. 5244657
281	36	21.1	459	1	US-08-087-732-2	Sequence 2, Appl	354	35	20.5	120	5	5433945-8	Patent No. 5433945
282	36	21.1	459	5	5210189-2	Patent No. 5210189	355	35	20.5	121	5	5223610-8	Patent No. 5223610
283	36	21.1	462	1	US-08-299-351-1	Sequence 1, Appl	356	35	20.5	138	2	US-08-888-497-32	Sequence 32, Appl
284	36	21.1	462	2	US-08-371-377-18	Sequence 18, Appl	357	35	20.5	138	4	PCT-US94-07926-32	Sequence 32, Appl
285	36	21.1	462	5	5225348-1	Patent No. 5225348	358	35	20.5	155	1	US-08-494-577-6	Sequence 6, Appl
286	36	21.1	476	1	US-08-565-655-6	Sequence 6, Appl	359	35	20.5	155	1	US-08-494-577-7	Sequence 7, Appl
287	36	21.1	509	1	US-07-779-890-6	Sequence 6, Appl	360	35	20.5	155	2	US-08-795-868-6	Sequence 6, Appl
288	36	21.1	509	2	US-07-779-890-6	Sequence 6, Appl	361	35	20.5	155	2	US-08-795-868-7	Sequence 7, Appl
289	36	21.1	509	1	US-09-008-962-3	Sequence 3, Appl	362	35	20.5	191	2	US-08-290-665A-173	Sequence 173, App
290	36	21.1	509	2	US-08-675-507-3	Sequence 3, Appl	363	35	20.5	191	4	PCT-US95-10398-173	Sequence 173, App
291	36	21.1	509	4	PCT-US93-05640-6	Sequence 6, Appl	364	35	20.5	192	1	US-08-086-428B-60	Sequence 60, Appl
292	36	21.1	511	1	US-08-278-635B-8	Sequence 8, Appl	365	35	20.5	192	1	US-08-086-428B-61	Sequence 61, Appl
293	36	21.1	511	3	US-08-464-258B-8	Sequence 8, Appl	366	35	20.5	192	1	US-08-086-428B-66	Sequence 66, Appl
294	36	21.1	532	3	US-08-481-435-12	Sequence 12, Appl	367	35	20.5	192	1	US-08-086-428B-73	Sequence 73, Appl
295	36	21.1	533	1	US-07-952-800-2	Sequence 2, Appl	368	35	20.5	192	1	US-08-086-428B-74	Sequence 74, Appl
296	36	21.1	536	1	US-08-164-614A-12	Sequence 12, Appl	369	35	20.5	192	2	US-08-468-570-60	Sequence 60, Appl
297	36	21.1	536	2	US-08-456-489B-12	Sequence 12, Appl	370	35	20.5	192	2	US-08-468-570-61	Sequence 61, Appl
298	36	21.1	553	3	US-08-481-435-11	Sequence 11, Appl	371	35	20.5	192	2	US-08-468-570-66	Sequence 66, Appl
299	36	21.1	663	1	US-07-912-015-2	Sequence 2, Appl	372	35	20.5	192	2	US-08-468-570-73	Sequence 73, Appl
300	36	21.1	823	3	US-08-481-435-4	Sequence 4, Appl	373	35	20.5	192	2	US-08-468-570-74	Sequence 74, Appl
301	36	21.1	846	3	US-08-481-435-9	Sequence 9, Appl	374	35	20.5	192	2	US-08-290-665A-60	Sequence 60, Appl
302	36	21.1	846	3	US-08-481-435-7	Sequence 7, Appl	375	35	20.5	192	2	US-08-290-665A-61	Sequence 61, Appl
303	36	21.1	844	3	US-08-481-435-8	Sequence 8, Appl	376	35	20.5	192	2	US-08-290-665A-66	Sequence 66, Appl
304	36	21.1	1153	1	US-08-314-917-2	Sequence 2, Appl	377	35	20.5	192	2	US-08-290-665A-73	Sequence 73, Appl
305	36	21.1	1153	1	US-08-265-046-2	Sequence 2, Appl	378	35	20.5	192	2	PCT-US95-10398-60	Sequence 60, Appl
306	36	21.1	1153	2	US-08-465-522-2	Sequence 2, Appl	379	35	20.5	192	4	PCT-US95-10398-74	Sequence 74, Appl
307	36	21.1	1153	4	PCT-US93-11401-2	Sequence 2, Appl	380	35	20.5	192	4	PCT-US95-10398-61	Sequence 61, Appl
308	36	21.1	1153	4	PCT-US95-07849-2	Sequence 2, Appl	381	35	20.5	192	4	PCT-US95-10398-66	Sequence 66, Appl
309	36	21.1	1367	2	US-08-249-687C-2	Sequence 2, Appl	382	35	20.5	192	4	PCT-US95-10398-73	Sequence 73, Appl
310	36	21.1	1367	2	US-08-625-819-2	Sequence 2, Appl	383	35	20.5	192	4	PCT-US95-10398-74	Sequence 74, Appl
311	36	21.1	1674	2	US-08-968-542C-12	Sequence 12, Appl	384	35	20.5	244	1	US-08-361-395-1	Sequence 1, Appl
312	36	21.1	2050	2	US-08-347-594A-2	Sequence 2, Appl	385	35	20.5	262	3	US-08-436-748-7	Sequence 7, Appl
313	36	21.1	2629	2	US-08-751-189-4	Sequence 4, Appl	386	35	20.5	262	3	US-08-483-857-7	Sequence 7, Appl
314	36	21.1	2629	2	US-09-060-836-4	Sequence 4, Appl	387	35	20.5	279	4	PCT-US91-00899-8	Sequence 8, Appl
315	35.5	20.8	125	1	US-08-170-360-5	Sequence 5, Appl	388	35	20.5	281	1	US-08-467-155A-7	Sequence 7, Appl
316	35.5	20.8	127	2	US-08-853-659A-56	Sequence 56, Appl	389	35	20.5	281	2	US-08-628-198-7	Sequence 7, Appl
317	35.5	20.8	128	3	US-08-804-439A-97	Sequence 97, Appl	390	35	20.5	281	4	PCT-US96-07343-2	Sequence 2, Appl
318	35.5	20.8	128	3	US-08-720-229-97	Sequence 97, Appl	391	35	20.5	352	1	US-08-482-577B-2	Sequence 2, Appl
319	35.5	20.8	191	2	US-08-290-665A-181	Sequence 181, App	392	35	20.5	364	1	US-08-680-726A-56	Sequence 56, Appl
320	35.5	20.8	191	2	US-08-290-665A-184	Sequence 184, App	393	35	20.5	368	1	US-08-093-372-4	Sequence 4, Appl

394	35	20.5	371	2	US-08-378-617A-12	Sequence 12, Appl	467	35	20.5	929	2	US-08-486-273A-40	Sequence 40, Appl
395	35	20.5	393	5	5182195-10	Patent No. 5182195	468	35	20.5	929	3	US-08-480-474-40	Sequence 40, Appl
396	35	20.5	394	1	US-07-914-281-4	Sequence 4, Appl	469	35	20.5	938	2	US-08-231-193A-2	Sequence 2, Appl
397	35	20.5	394	1	US-08-332-246-4	Sequence 4, Appl	470	35	20.5	938	2	US-08-486-273A-2	Sequence 2, Appl
398	35	20.5	394	1	US-08-525-058A-4	Sequence 4, Appl	471	35	20.5	938	3	US-08-480-474-2	Sequence 2, Appl
399	35	20.5	394	2	US-08-696-731-4	Sequence 4, Appl	472	35	20.5	943	2	US-08-231-193A-36	Sequence 36, Appl
400	35	20.5	394	4	PCT-US91-00899-9	Sequence 9, Appl	473	35	20.5	943	2	US-08-486-273A-36	Sequence 36, Appl
401	35	20.5	400	2	US-08-733-825-2	Sequence 2, Appl	474	35	20.5	943	3	US-08-480-474-36	Sequence 36, Appl
402	35	20.5	426	2	US-08-560-398-4	Sequence 4, Appl	475	35	20.5	947	2	US-08-887-518-2	Sequence 2, Appl
403	35	20.5	431	1	US-08-381-433A-2	Sequence 2, Appl	476	35	20.5	947	2	US-09-023-321-2	Sequence 2, Appl
404	35	20.5	434	2	US-08-815-718-3	Sequence 3, Appl	477	35	20.5	947	2	US-09-032-475-2	Sequence 2, Appl
405	35	20.5	446	1	US-08-833-610-5	Sequence 5, Appl	478	35	20.5	959	2	US-08-231-193A-24	Sequence 24, Appl
406	35	20.5	448	1	US-08-366-779-5	Sequence 5, Appl	479	35	20.5	959	3	US-08-486-273A-24	Sequence 24, Appl
407	35	20.5	448	1	US-08-789-936-5	Sequence 5, Appl	480	35	20.5	959	3	US-08-480-474-24	Sequence 24, Appl
408	35	20.5	456	1	US-08-435-933-6	Sequence 6, Appl	481	35	20.5	976	2	US-08-231-193A-22	Sequence 22, Appl
409	35	20.5	456	1	PCT-US96-08035-6	Sequence 6, Appl	482	35	20.5	976	2	US-08-486-273A-22	Sequence 22, Appl
410	35	20.5	483	2	US-08-583-562B-12	Sequence 12, Appl	483	35	20.5	976	3	US-08-480-474-22	Sequence 22, Appl
411	35	20.5	483	2	US-08-779-113-12	Sequence 12, Appl	484	35	20.5	995	4	PCT-US95-04910-14	Sequence 14, Appl
412	35	20.5	507	1	US-08-363-475-19	Sequence 19, Appl	485	35	20.5	997	2	US-08-231-193A-38	Sequence 38, Appl
413	35	20.5	532	1	US-08-363-475-22	Sequence 22, Appl	486	35	20.5	997	2	US-08-486-273A-38	Sequence 38, Appl
414	35	20.5	533	1	US-08-294-872-2	Sequence 2, Appl	487	35	20.5	997	3	US-08-480-474-38	Sequence 38, Appl
415	35	20.5	533	4	PCT-US95-09823-2	Sequence 2, Appl	488	35	20.5	1311	1	US-08-340-011-5	Sequence 5, Appl
416	35	20.5	559	2	US-08-884-072-6	Sequence 6, Appl	489	35	20.5	1338	3	US-08-750-141A-3	Sequence 3, Appl
417	35	20.5	566	2	US-07-862-588B-4	Sequence 4, Appl	490	35	20.5	1940	2	US-08-644-271-30	Sequence 30, Appl
418	35	20.5	582	1	US-08-261-086-8	Sequence 8, Appl	491	35	20.5	3038	1	US-08-450-332-2	Sequence 2, Appl
419	35	20.5	610	1	US-08-410-804-1	Sequence 1, Appl	492	35	20.5	3038	2	US-08-637-640-2	Sequence 2, Appl
420	35	20.5	610	1	US-08-259-514-1	Sequence 1, Appl	493	35	20.5	3079	4	PCT-US94-00198-4	Sequence 4, Appl
421	35	20.5	610	2	US-08-858-311-1	Sequence 1, Appl	494	34.5	20.2	38	1	US-08-243-728-56	Sequence 56, Appl
422	35	20.5	610	2	US-08-799-138-6	Sequence 6, Appl	495	34.5	20.2	87	1	US-08-271-354-6	Sequence 6, Appl
423	35	20.5	617	4	PCT-US92-06965A-35	Sequence 35, Appl	496	34.5	20.2	87	4	US-08-565-861-6	Sequence 6, Appl
424	35	20.5	638	2	US-08-681-151-3	Sequence 3, Appl	497	34.5	20.2	106	1	PCT-US94-07658-6	Sequence 6, Appl
425	35	20.5	661	1	US-08-232-538-12	Sequence 12, Appl	498	34.5	20.2	159	1	US-08-491-976-1	Sequence 1, Appl
426	35	20.5	661	2	US-08-786-164-12	Sequence 12, Appl	499	34.5	20.2	159	1	US-08-491-976-3	Sequence 3, Appl
427	35	20.5	687	1	US-08-232-538-6	Sequence 6, Appl	500	34.5	20.2	192	1	US-08-086-428B-94	Sequence 94, Appl
428	35	20.5	687	2	US-08-786-164-6	Sequence 6, Appl	501	34.5	20.2	192	2	US-08-468-570-94	Sequence 94, Appl
429	35	20.5	695	5	5460961-5	Patent No. 5460961	502	34.5	20.2	192	4	PCT-US95-10398-94	Sequence 94, Appl
430	35	20.5	709	3	US-08-968-563-18	Sequence 18, Appl	503	34.5	20.2	192	4	PCT-US95-10398-94	Sequence 94, Appl
431	35	20.5	712	1	US-08-587-889-2	Sequence 2, Appl	504	34.5	20.2	274	2	US-08-469-537A-55	Sequence 55, Appl
432	35	20.5	712	2	US-08-980-060-5	Sequence 5, Appl	505	34.5	20.2	274	2	US-08-857-224B-71	Sequence 71, Appl
433	35	20.5	712	4	PCT-US96-09193-2	Sequence 2, Appl	506	34.5	20.2	293	2	US-08-701-191A-26	Sequence 26, Appl
434	35	20.5	758	1	US-08-289-112-2	Sequence 2, Appl	507	34.5	20.2	314	2	US-08-928-615-2	Sequence 2, Appl
435	35	20.5	758	2	US-08-874-678-1	Sequence 1, Appl	508	34.5	20.2	320	1	US-08-565-386-17	Sequence 17, Appl
436	35	20.5	777	2	US-08-231-193A-16	Sequence 16, Appl	509	34.5	20.2	328	3	US-09-225-244-2	Sequence 2, Appl
437	35	20.5	777	2	US-08-486-273A-16	Sequence 16, Appl	510	34.5	20.2	344	2	US-09-055-087-3	Sequence 3, Appl
438	35	20.5	777	3	US-08-480-474-16	Sequence 16, Appl	511	34.5	20.2	361	1	US-08-415-751-7	Sequence 7, Appl
439	35	20.5	780	1	US-08-232-538-14	Sequence 14, Appl	512	34.5	20.2	365	2	US-08-724-974A-3	Sequence 3, Appl
440	35	20.5	780	2	US-08-786-164-14	Sequence 14, Appl	513	34.5	20.2	382	3	US-08-811-177A-2	Sequence 2, Appl
441	35	20.5	788	2	US-08-918-914-4	Sequence 4, Appl	514	34.5	20.2	470	2	US-08-691-814B-2	Sequence 2, Appl
442	35	20.5	854	2	US-08-231-193A-32	Sequence 32, Appl	515	34.5	20.2	531	2	US-08-724-974A-2	Sequence 2, Appl
443	35	20.5	854	2	US-08-486-273A-32	Sequence 32, Appl	516	34.5	20.2	592	2	US-08-736-770-6	Sequence 6, Appl
444	35	20.5	854	3	US-08-480-474-32	Sequence 32, Appl	517	34.5	20.2	617	1	US-08-279-700-4	Sequence 4, Appl
445	35	20.5	857	2	US-08-779-113-2	Sequence 2, Appl	518	34.5	20.2	617	1	US-08-279-700-6	Sequence 6, Appl
446	35	20.5	870	2	US-08-231-193A-30	Sequence 30, Appl	519	34.5	20.2	617	1	US-08-279-700-8	Sequence 8, Appl
447	35	20.5	870	2	US-08-486-273A-30	Sequence 30, Appl	520	34.5	20.2	617	1	US-08-279-700-10	Sequence 10, Appl
448	35	20.5	870	3	US-08-480-474-30	Sequence 30, Appl	521	34.5	20.2	617	1	US-08-279-700-12	Sequence 12, Appl
449	35	20.5	875	2	US-08-231-193A-28	Sequence 28, Appl	522	34.5	20.2	617	1	US-08-279-700-21	Sequence 21, Appl
450	35	20.5	875	2	US-08-486-273A-28	Sequence 28, Appl	523	34.5	20.2	898	1	US-08-465-995A-2	Sequence 2, Appl
451	35	20.5	875	3	US-08-480-474-28	Sequence 28, Appl	524	34.5	20.2	898	1	US-08-465-995A-4	Sequence 4, Appl
452	35	20.5	891	2	US-08-231-193A-26	Sequence 26, Appl	525	34.5	20.2	898	2	US-08-465-994C-2	Sequence 2, Appl
453	35	20.5	891	3	US-08-486-273A-26	Sequence 26, Appl	526	34.5	20.2	898	2	US-08-465-994C-4	Sequence 4, Appl
454	35	20.5	891	3	US-08-480-474-26	Sequence 26, Appl	527	34.5	20.2	898	2	US-08-966-145-2	Sequence 2, Appl
455	35	20.5	908	2	US-08-231-193A-34	Sequence 34, Appl	528	34.5	20.2	898	2	US-08-966-145-4	Sequence 4, Appl
456	35	20.5	908	2	US-08-486-273A-34	Sequence 34, Appl	529	34.5	20.2	920	1	US-08-101-593-2	Sequence 2, Appl
457	35	20.5	908	3	US-08-480-474-34	Sequence 34, Appl	530	34.5	20.2	920	1	US-08-101-593-4	Sequence 4, Appl
458	35	20.5	920	1	US-08-026-138E-9	Sequence 9, Appl	531	34.5	20.2	1648	2	US-08-662-227-35	Sequence 35, Appl
459	35	20.5	920	1	US-08-026-138E-10	Sequence 10, Appl	532	34	19.9	12	5	5223481-2	Patent No. 5223481
460	35	20.5	922	2	US-08-231-193A-14	Sequence 14, Appl	533	34	19.9	12	5	5223481-3	Patent No. 5223481
461	35	20.5	922	2	US-08-231-193A-20	Sequence 20, Appl	534	34	19.9	12	5	5223481-4	Patent No. 5223481
462	35	20.5	922	2	US-08-486-273A-14	Sequence 14, Appl	535	34	19.9	12	5	5455227-1	Patent No. 5455227
463	35	20.5	922	2	US-08-486-273A-20	Sequence 20, Appl	536	34	19.9	12	5	5455227-3	Patent No. 5455227
464	35	20.5	922	3	US-08-480-474-14	Sequence 14, Appl	537	34	19.9	12	5	5455227-6	Patent No. 5455227
465	35	20.5	922	3	US-08-480-474-20	Sequence 20, Appl	538	34	19.9	15	4	PCT-US93-11703-29	Sequence 29, Appl
466	35	20.5	929	2	US-08-231-193A-40	Sequence 40, Appl	539	34	19.9	17	5	5223481-8	Patent No. 5223481

540	34	19.9	17	5	5455227-4	Patent No. 5455227	613	34	19.9	370	1	US-07-966-187-26	Sequence 26, Appl
541	34	19.9	25	1	US-08-281-702A-4	Sequence 4, Appli	614	34	19.9	373	2	US-08-559-524A-4	Sequence 4, Appli
542	34	19.9	25	2	US-08-618-917-4	Sequence 4, Appli	615	34	19.9	383	5	5470718-5	Patent No. 5470718
543	34	19.9	35	1	US-08-418-893D-16	Sequence 16, Appl	616	34	19.9	388	1	US-08-499-568-11	Sequence 11, Appl
544	34	19.9	35	1	US-08-343-427B-11	Sequence 11, Appl	617	34	19.9	388	2	US-08-793-958-11	Sequence 11, Appl
545	34	19.9	35	2	US-08-652-450A-11	Sequence 11, Appl	618	34	19.9	389	2	US-08-846-762-82	Sequence 82, Appl
546	34	19.9	35	2	US-08-652-450A-5	Sequence 5, Appli	619	34	19.9	393	1	US-08-220-151-23	Sequence 23, Appl
547	34	19.9	35	5	5223481-5	Patent No. 5223481	620	34	19.9	393	1	US-08-499-568-15	Sequence 15, Appl
548	34	19.9	35	5	5223481-6	Patent No. 5223481	621	34	19.9	393	1	US-08-413-118-23	Sequence 23, Appl
549	34	19.9	35	5	5223481-7	Patent No. 5223481	622	34	19.9	393	2	US-08-793-958-15	Sequence 15, Appl
550	34	19.9	35	5	5455227-2	Patent No. 5455227	623	34	19.9	393	2	US-08-956-998-2	Sequence 2, Appli
551	34	19.9	38	1	US-08-243-728-63	Sequence 63, Appl	624	34	19.9	393	3	US-08-473-446-23	Sequence 23, Appl
552	34	19.9	38	1	US-08-243-728-64	Sequence 64, Appl	625	34	19.9	394	3	US-08-499-568-4	Sequence 4, Appli
553	34	19.9	62	1	US-07-662-005A-14	Sequence 14, Appl	626	34	19.9	394	2	US-08-793-958-4	Sequence 4, Appli
554	34	19.9	99	2	US-09-047-125-21	Sequence 21, Appl	627	34	19.9	414	1	US-08-255-471-9	Sequence 9, Appli
555	34	19.9	124	1	US-08-170-360-4	Sequence 4, Appli	628	34	19.9	416	2	US-08-698-407-5	Sequence 5, Appli
556	34	19.9	124	2	US-08-888-497-39	Sequence 39, Appl	629	34	19.9	416	3	US-09-195-855-5	Sequence 5, Appli
557	34	19.9	124	4	PCT-US94-07926-39	Sequence 39, Appl	630	34	19.9	438	1	US-08-097-827-11	Sequence 11, Appl
558	34	19.9	134	1	US-08-482-728A-13	Sequence 13, Appl	631	34	19.9	438	1	US-08-494-574-11	Sequence 11, Appl
559	34	19.9	144	2	US-08-186-895-10	Sequence 10, Appl	632	34	19.9	441	1	US-08-403-866-10	Sequence 10, Appl
560	34	19.9	144	2	US-08-888-497-37	Sequence 37, Appl	633	34	19.9	442	3	US-08-993-359-28	Sequence 28, Appl
561	34	19.9	144	4	PCT-US94-07926-37	Sequence 37, Appl	634	34	19.9	452	1	US-08-117-361C-1	Sequence 1, Appli
562	34	19.9	160	2	US-08-162-402B-11	Sequence 11, Appl	635	34	19.9	462	1	US-08-458-023B-2	Sequence 2, Appli
563	34	19.9	164	1	US-08-808-303-8	Sequence 8, Appli	636	34	19.9	463	2	US-08-162-402B-9	Sequence 9, Appli
564	34	19.9	164	1	US-08-808-303-12	Sequence 12, Appl	637	34	19.9	463	2	US-09-111-556A-2	Sequence 2, Appli
565	34	19.9	175	1	US-08-145-995A-5	Sequence 5, Appli	638	34	19.9	475	2	US-08-698-407-4	Sequence 4, Appli
566	34	19.9	175	2	US-08-451-747-5	Sequence 5, Appli	639	34	19.9	475	3	US-09-195-855-4	Sequence 4, Appli
567	34	19.9	192	1	US-08-086-428B-63	Sequence 63, Appl	640	34	19.9	477	1	US-07-735-065-2	Sequence 2, Appli
568	34	19.9	192	1	US-08-086-428B-65	Sequence 65, Appl	641	34	19.9	477	1	US-08-469-202-12	Sequence 12, Appl
569	34	19.9	192	1	US-08-468-570-63	Sequence 63, Appl	642	34	19.9	477	2	US-08-484-434C-12	Sequence 12, Appl
570	34	19.9	192	2	US-08-468-570-65	Sequence 65, Appl	643	34	19.9	490	1	US-08-361-611-2	Sequence 2, Appli
571	34	19.9	192	2	US-08-290-665A-63	Sequence 63, Appl	644	34	19.9	490	1	US-08-565-655-2	Sequence 2, Appli
572	34	19.9	192	2	US-08-290-665A-65	Sequence 65, Appl	645	34	19.9	490	2	US-08-946-967-2	Sequence 2, Appli
573	34	19.9	192	4	PCT-US95-10398-63	Sequence 63, Appl	646	34	19.9	497	1	US-08-252-492-2	Sequence 2, Appli
574	34	19.9	192	4	PCT-US95-10398-65	Sequence 65, Appl	647	34	19.9	497	2	US-08-727-126-2	Sequence 2, Appli
575	34	19.9	204	1	US-08-652-859-2	Sequence 2, Appli	648	34	19.9	514	1	US-08-361-920-21	Sequence 21, Appl
576	34	19.9	204	2	US-08-919-706-2	Sequence 2, Appli	649	34	19.9	514	1	US-08-479-939-21	Sequence 21, Appl
577	34	19.9	204	2	US-09-153-751-2	Sequence 2, Appli	650	34	19.9	514	1	US-08-483-432-21	Sequence 21, Appl
578	34	19.9	207	2	US-08-531-525-35	Sequence 35, Appl	651	34	19.9	552	4	PCT-US93-07832-23	Sequence 23, Appl
579	34	19.9	207	2	US-08-718-270A-35	Sequence 35, Appl	652	34	19.9	553	1	US-08-565-386-12	Sequence 12, Appl
580	34	19.9	211	1	US-08-631-607-4	Sequence 4, Appli	653	34	19.9	592	2	US-08-599-171A-30	Sequence 30, Appl
581	34	19.9	215	1	US-08-431-080-22	Sequence 22, Appl	654	34	19.9	592	2	US-08-646-590B-30	Sequence 30, Appl
582	34	19.9	215	2	US-08-938-534-22	Sequence 22, Appl	655	34	19.9	592	3	US-09-049-226-30	Sequence 30, Appl
583	34	19.9	262	2	US-08-038-761A-1	Sequence 1, Appli	656	34	19.9	596	2	US-08-392-806A-2	Sequence 2, Appli
584	34	19.9	263	1	US-07-662-005A-12	Sequence 12, Appl	657	34	19.9	600	2	US-08-392-806A-4	Sequence 4, Appli
585	34	19.9	290	1	US-08-202-186-26	Sequence 26, Appl	658	34	19.9	606	2	US-08-392-806A-6	Sequence 6, Appli
586	34	19.9	302	2	US-09-031-485-38	Sequence 38, Appl	659	34	19.9	607	1	US-07-959-943-7	Sequence 7, Appli
587	34	19.9	302	2	US-08-847-429A-38	Sequence 38, Appl	660	34	19.9	607	1	US-07-879-617A-12	Sequence 12, Appl
588	34	19.9	303	2	US-09-031-485-23	Sequence 23, Appl	661	34	19.9	607	1	US-08-753-985-12	Sequence 12, Appl
589	34	19.9	303	2	US-08-847-429A-23	Sequence 23, Appl	662	34	19.9	610	3	US-08-484-661A-26	Sequence 26, Appl
590	34	19.9	308	1	US-08-499-568-2	Sequence 2, Appli	663	34	19.9	610	4	PCT-US96-09641-26	Sequence 26, Appl
591	34	19.9	308	2	US-08-793-958-2	Sequence 2, Appli	664	34	19.9	630	1	US-07-959-943-9	Sequence 9, Appli
592	34	19.9	308	2	US-08-164-292B-18	Sequence 18, Appl	665	34	19.9	653	2	US-07-782-298-2	Sequence 2, Appli
593	34	19.9	308	3	US-08-845-623-18	Sequence 18, Appl	666	34	19.9	703	2	US-08-835-099A-1	Sequence 1, Appli
594	34	19.9	320	2	US-08-480-229C-20	Sequence 20, Appl	667	34	19.9	712	2	US-08-835-099A-2	Sequence 2, Appli
595	34	19.9	320	2	US-08-659-235C-20	Sequence 20, Appl	668	34	19.9	879	1	US-08-306-546C-2	Sequence 2, Appli
596	34	19.9	359	1	US-08-181-271A-98	Sequence 98, Appl	669	34	19.9	879	2	US-08-530-524A-2	Sequence 2, Appli
597	34	19.9	359	1	US-08-449-315-98	Sequence 98, Appl	670	34	19.9	899	1	US-08-365-689-2	Sequence 2, Appli
598	34	19.9	359	1	US-08-444-803-98	Sequence 98, Appl	671	34	19.9	899	1	US-08-145-138A-2	Sequence 2, Appli
599	34	19.9	359	1	US-08-047-413-15	Sequence 15, Appl	672	34	19.9	933	1	US-07-747-781-2	Sequence 2, Appli
600	34	19.9	359	1	US-08-449-043-98	Sequence 98, Appl	673	34	19.9	933	2	PCT-US92-08888-2	Sequence 2, Appli
601	34	19.9	359	1	US-08-456-265A-98	Sequence 98, Appl	674	34	19.9	973	4	US-08-683-262B-75	Sequence 75, Appl
602	34	19.9	359	1	US-08-455-416-98	Sequence 98, Appl	675	34	19.9	1290	1	US-08-138-641-2	Sequence 2, Appli
603	34	19.9	359	1	US-08-455-244-98	Sequence 98, Appl	676	34	19.9	1290	1	US-08-138-133-2	Sequence 2, Appli
604	34	19.9	359	1	US-08-454-876-98	Sequence 98, Appl	677	34	19.9	1514	2	US-08-853-310-4	Sequence 4, Appli
605	34	19.9	359	2	US-08-457-364-98	Sequence 98, Appl	678	34	19.9	1713	2	PCT-US94-10261A-24	Sequence 24, Appl
606	34	19.9	359	2	US-08-456-262-98	Sequence 98, Appl	679	34	19.9	1745	4	US-09-031-485-33	Sequence 33, Appl
607	34	19.9	359	2	US-08-456-240-98	Sequence 98, Appl	680	34	19.9	1745	3	US-08-847-429A-33	Sequence 33, Appl
608	34	19.9	359	2	US-08-455-736-98	Sequence 98, Appl	681	34	19.9	2813	3	US-08-896-449A-2	Sequence 2, Appli
609	34	19.9	359	2	US-08-971-217-98	Sequence 98, Appl	682	34	19.9	3011	1	US-08-188-281B-1	Sequence 1, Appli
610	34	19.9	362	2	US-08-185-828A-13	Sequence 13, Appl	683	34	19.9	3011	1	US-08-453-552-1	Sequence 1, Appli
611	34	19.9	368	4	PCT-US93-11703-24	Sequence 24, Appl	684	34	19.9	3011	2	US-08-710-637-1	Sequence 1, Appli
612	34	19.9	369	2	US-08-139-609-1	Sequence 1, Appli	685	34	19.9	3011	4	PCT-US93-00907-1	Sequence 1, Appli

686	34	19.9	3011	4	PCT-US94-07280-1	Sequence 1, Appli	759	33	19.3	100	2	US-09-047-125-27	Sequence 27, Appl
687	34	19.9	3011	4	PCT-US95-01087-1	Sequence 1, Appli	760	33	19.3	104	2	US-08-727-688-32	Sequence 32, Appl
688	34	19.9	3119	1	US-08-246-982A-16	Sequence 16, Appl	761	33	19.3	116	1	US-08-702-344-16	Sequence 16, Appl
689	34	19.9	3119	1	US-08-433-265-16	Sequence 16, Appl	762	33	19.3	120	1	US-07-923-206-2	Sequence 2, Appli
690	34	19.9	3457	1	US-08-416-603-4	Sequence 4, Appli	763	33	19.3	120	2	US-08-313-185-43	Sequence 43, Appl
691	33.5	19.6	39	1	US-08-243-728-58	Sequence 58, Appl	764	33	19.3	120	2	US-08-459-499-2	Sequence 2, Appli
692	33.5	19.6	76	2	US-08-248-839C-182	Sequence 182, App	765	33	19.3	127	1	US-08-455-550-11	Sequence 11, Appl
693	33.5	19.6	101	2	US-08-466-860-75	Sequence 75, Appl	766	33	19.3	137	1	US-08-676-169-2	Sequence 2, Appli
694	33.5	19.6	113	2	US-08-466-860-8	Sequence 9, Appli	767	33	19.3	138	2	US-08-757-036-1	Sequence 1, Appli
695	33.5	19.6	145	2	US-08-640-977-6	Sequence 29, Appl	768	33	19.3	148	1	US-08-207-904-15	Sequence 15, Appl
696	33.5	19.6	168	2	US-08-702-105A-29	Sequence 29, Appl	769	33	19.3	163	3	US-09-053-197A-22	Sequence 22, Appl
697	33.5	19.6	168	2	US-08-702-105A-31	Sequence 31, Appl	770	33	19.3	166	2	US-08-483-695-3	Sequence 3, Appli
698	33.5	19.6	168	3	US-08-702-110A-29	Sequence 29, Appl	771	33	19.3	166	2	US-08-483-695-31	Sequence 31, Appl
699	33.5	19.6	168	3	US-08-702-110A-31	Sequence 31, Appl	772	33	19.3	166	2	US-08-483-695-32	Sequence 32, Appl
700	33.5	19.6	207	2	US-08-640-977-5	Sequence 5, Appli	773	33	19.3	166	2	US-08-483-695-33	Sequence 33, Appl
701	33.5	19.6	263	2	US-08-333-476-67	Sequence 67, Appl	774	33	19.3	166	2	US-08-483-695-34	Sequence 34, Appl
702	33.5	19.6	266	2	US-08-640-977-4	Sequence 4, Appli	775	33	19.3	166	2	US-07-965-285-3	Sequence 3, Appli
703	33.5	19.6	326	2	US-08-640-977-1	Sequence 1, Appli	776	33	19.3	166	2	US-07-965-285-31	Sequence 31, Appl
704	33.5	19.6	338	2	US-09-004-502-5	Sequence 5, Appli	777	33	19.3	166	2	US-07-965-285-32	Sequence 32, Appl
705	33.5	19.6	390	1	US-07-817-920-6	Sequence 6, Appli	778	33	19.3	166	2	US-07-965-285-33	Sequence 33, Appl
706	33.5	19.6	390	1	US-08-117-006-6	Sequence 6, Appli	779	33	19.3	166	2	US-07-965-285-34	Sequence 34, Appl
707	33.5	19.6	390	1	US-08-216-594-6	Sequence 6, Appli	780	33	19.3	166	2	US-08-487-231-3	Sequence 3, Appli
708	33.5	19.6	390	2	US-08-461-812-4	Sequence 4, Appli	781	33	19.3	166	2	US-08-487-231-31	Sequence 31, Appl
709	33.5	19.6	390	2	US-08-157-185-15	Sequence 15, Appl	782	33	19.3	166	2	US-08-487-231-32	Sequence 32, Appl
710	33.5	19.6	390	4	PCT-US93-00149-6	Sequence 6, Appli	783	33	19.3	166	2	US-08-487-231-33	Sequence 33, Appl
711	33.5	19.6	398	1	US-08-370-542-6	Sequence 6, Appli	784	33	19.3	166	2	US-08-487-231-34	Sequence 34, Appl
712	33.5	19.6	398	1	US-08-542-358-6	Sequence 6, Appli	785	33	19.3	168	3	US-09-053-197A-20	Sequence 20, Appl
713	33.5	19.6	520	2	US-08-482-728A-4	Sequence 4, Appli	786	33	19.3	182	3	US-09-261-471-5	Sequence 5, Appli
714	33.5	19.6	550	4	PCT-US93-03027-1	Sequence 1, Appli	787	33	19.3	190	1	US-07-681-701-16	Sequence 16, Appl
715	33.5	19.6	555	2	US-08-780-835B-2	Sequence 2, Appli	788	33	19.3	191	2	US-08-290-665A-155	Sequence 155, App
716	33.5	19.6	601	2	US-08-606-288-7	Sequence 7, Appli	789	33	19.3	191	2	US-08-290-665A-156	Sequence 156, App
717	33.5	19.6	601	2	US-08-606-288-10	Sequence 10, Appli	790	33	19.3	191	2	US-08-290-665A-157	Sequence 157, App
718	33.5	19.6	617	1	US-08-279-700-2	Sequence 2, Appli	791	33	19.3	191	2	US-08-290-665A-158	Sequence 158, App
719	33.5	19.6	617	1	US-08-348-891A-6	Sequence 6, Appli	792	33	19.3	191	2	US-08-290-665A-159	Sequence 159, App
720	33.5	19.6	617	2	US-08-905-817-6	Sequence 6, Appli	793	33	19.3	191	2	US-08-290-665A-160	Sequence 160, App
721	33.5	19.6	725	1	US-08-448-170-4	Sequence 4, Appli	794	33	19.3	191	2	US-08-290-665A-162	Sequence 162, App
722	33.5	19.6	773	2	US-08-484-101B-42	Sequence 42, Appli	795	33	19.3	191	2	US-08-290-665A-163	Sequence 163, App
723	33.5	19.6	773	2	US-08-484-101B-44	Sequence 44, Appli	796	33	19.3	191	2	US-08-290-665A-164	Sequence 164, App
724	33.5	19.6	861	1	US-08-484-105-18	Sequence 18, Appl	797	33	19.3	191	2	US-08-290-665A-165	Sequence 165, App
725	33.5	19.6	861	1	US-08-484-106-18	Sequence 18, Appl	798	33	19.3	191	2	US-08-290-665A-166	Sequence 166, App
726	33.5	19.6	942	2	US-08-884-681-4	Sequence 4, Appli	799	33	19.3	191	2	US-08-290-665A-167	Sequence 167, App
727	33.5	19.6	942	3	US-08-560-005-6	Sequence 6, Appli	800	33	19.3	191	2	US-08-290-665A-168	Sequence 168, App
728	33.5	19.6	1055	3	US-09-031-563-27	Sequence 27, Appl	801	33	19.3	191	2	US-08-290-665A-169	Sequence 169, App
729	33.5	19.6	1088	3	US-08-633-768A-1	Sequence 1, Appli	802	33	19.3	191	2	US-08-290-665A-170	Sequence 170, App
730	33.5	19.6	1162	2	US-08-683-566A-15	Sequence 15, Appl	803	33	19.3	191	2	US-08-290-665A-171	Sequence 171, App
731	33.5	19.6	1162	2	US-08-023-610-15	Sequence 15, Appl	804	33	19.3	191	2	US-08-290-665A-172	Sequence 172, App
732	33.5	19.6	1162	2	US-08-288-065A-15	Sequence 15, Appl	805	33	19.3	191	2	US-08-290-665A-174	Sequence 174, App
733	33.5	19.6	1162	2	US-08-362-240A-15	Sequence 15, Appl	806	33	19.3	191	2	US-08-290-665A-175	Sequence 175, App
734	33.5	19.6	1162	4	PCT-US95-10245-15	Sequence 15, Appl	807	33	19.3	191	2	US-08-290-665A-176	Sequence 176, App
735	33.5	19.6	1315	3	US-09-031-563-2	Sequence 2, Appli	808	33	19.3	191	2	US-08-290-665A-177	Sequence 177, App
736	33.5	19.6	1315	3	US-08-031-563-25	Sequence 25, Appl	809	33	19.3	191	2	US-08-290-665A-178	Sequence 178, App
737	33	19.3	10	2	US-08-764-640-76	Sequence 76, Appl	810	33	19.3	191	2	US-08-290-665A-179	Sequence 179, App
738	33	19.3	38	1	US-08-802-981-100	Sequence 100, App	811	33	19.3	191	2	US-08-290-665A-180	Sequence 180, App
739	33	19.3	38	1	US-08-393-036-1	Sequence 1, Appli	812	33	19.3	191	2	US-08-290-665A-182	Sequence 182, App
740	33	19.3	38	1	US-08-288-395-1	Sequence 1, Appli	813	33	19.3	191	2	US-08-290-665A-183	Sequence 183, App
741	33	19.3	38	1	US-08-243-728-55	Sequence 55, Appli	814	33	19.3	191	2	US-08-290-665A-185	Sequence 185, App
742	33	19.3	38	4	PCT-US93-03795-4	Sequence 4, Appli	815	33	19.3	191	2	US-08-290-665A-186	Sequence 186, App
743	33	19.3	39	1	US-08-243-728-65	Sequence 65, Appl	816	33	19.3	191	2	US-08-290-665A-191	Sequence 191, App
744	33	19.3	39	1	US-08-243-728-66	Sequence 66, Appl	817	33	19.3	191	2	US-08-290-665A-192	Sequence 192, App
745	33	19.3	39	4	PCT-US93-03795-5	Sequence 5, Appli	818	33	19.3	191	2	US-08-290-665A-193	Sequence 193, App
746	33	19.3	50	1	US-08-171-383A-3	Sequence 3, Appli	819	33	19.3	191	2	US-08-290-665A-194	Sequence 194, App
747	33	19.3	50	1	US-08-461-041-3	Sequence 3, Appli	820	33	19.3	191	2	US-08-290-665A-195	Sequence 195, App
748	33	19.3	51	1	US-08-188-228-34	Sequence 34, Appl	821	33	19.3	191	2	US-08-290-665A-196	Sequence 196, App
749	33	19.3	51	1	US-08-332-643-34	Sequence 34, Appl	822	33	19.3	191	2	US-08-290-665A-197	Sequence 197, App
750	33	19.3	51	1	US-08-332-638-34	Sequence 34, Appl	823	33	19.3	191	2	US-08-290-665A-198	Sequence 198, App
751	33	19.3	54	3	US-08-301-162-12	Sequence 12, Appl	824	33	19.3	191	2	US-08-290-665A-199	Sequence 199, App
752	33	19.3	68	1	US-08-171-383A-4	Sequence 4, Appli	825	33	19.3	191	2	US-08-290-665A-200	Sequence 200, App
753	33	19.3	68	1	US-08-461-041-4	Sequence 4, Appli	826	33	19.3	191	2	US-08-290-665A-201	Sequence 201, App
754	33	19.3	76	1	US-08-324-977-18	Sequence 18, Appl	827	33	19.3	191	2	US-08-290-665A-202	Sequence 202, App
755	33	19.3	76	2	US-08-384-616-18	Sequence 18, Appl	828	33	19.3	191	2	US-08-290-665A-203	Sequence 203, App
756	33	19.3	76	2	US-08-904-686A-18	Sequence 18, Appl	829	33	19.3	191	2	US-08-290-665A-204	Sequence 204, App
757	33	19.3	94	1	US-08-150-203A-6	Sequence 6, Appli	830	33	19.3	191	2	US-08-290-665A-205	Sequence 205, App
758	33	19.3	94	1	US-08-434-730-6	Sequence 6, Appli	831	33	19.3	191	2	US-08-290-665A-206	Sequence 206, App

832	33	19.3	191	4	PCT-US95-10398-155	Sequence 155, App	905	33	19.3	232	1	US-07-990-301A-4	Sequence 4, Appli
833	33	19.3	191	4	PCT-US95-10398-156	Sequence 156, App	906	33	19.3	246	2	US-08-978-404B-44	Sequence 44, Appl
834	33	19.3	191	4	PCT-US95-10398-157	Sequence 157, App	907	33	19.3	256	2	US-08-799-454A-3	Sequence 3, Appli
835	33	19.3	191	4	PCT-US95-10398-158	Sequence 158, App	908	33	19.3	256	4	PCT-US95-09927-3	Sequence 3, Appli
836	33	19.3	191	4	PCT-US95-10398-159	Sequence 159, App	909	33	19.3	256	4	PCT-US95-10904-67	Sequence 67, Appl
837	33	19.3	191	4	PCT-US95-10398-160	Sequence 160, App	910	33	19.3	286	3	US-08-301-162-16	Sequence 16, Appl
838	33	19.3	191	4	PCT-US95-10398-161	Sequence 161, App	911	33	19.3	286	1	US-08-467-155A-9	Sequence 9, Appli
839	33	19.3	191	4	PCT-US95-10398-162	Sequence 162, App	912	33	19.3	286	2	US-08-628-198-9	Sequence 9, Appli
840	33	19.3	191	4	PCT-US95-10398-163	Sequence 163, App	913	33	19.3	286	4	PCT-US96-07343-9	Sequence 9, Appli
841	33	19.3	191	4	PCT-US95-10398-164	Sequence 164, App	914	33	19.3	290	1	US-08-457-245-7	Sequence 7, Appli
842	33	19.3	191	4	PCT-US95-10398-165	Sequence 165, App	915	33	19.3	307	1	US-08-713-828-1	Sequence 1, Appli
843	33	19.3	191	4	PCT-US95-10398-166	Sequence 166, App	916	33	19.3	307	2	US-08-919-627-1	Sequence 1, Appli
844	33	19.3	191	4	PCT-US95-10398-167	Sequence 167, App	917	33	19.3	307	2	US-09-096-245-1	Sequence 1, Appli
845	33	19.3	191	4	PCT-US95-10398-168	Sequence 168, App	918	33	19.3	336	2	US-08-784-651-2	Sequence 2, Appli
846	33	19.3	191	4	PCT-US95-10398-169	Sequence 169, App	919	33	19.3	340	1	US-08-462-195-2	Sequence 2, Appli
847	33	19.3	191	4	PCT-US95-10398-170	Sequence 170, App	920	33	19.3	340	1	US-08-462-195-4	Sequence 2, Appli
848	33	19.3	191	4	PCT-US95-10398-171	Sequence 171, App	921	33	19.3	340	2	US-08-636-883-2	Sequence 2, Appli
849	33	19.3	191	4	PCT-US95-10398-172	Sequence 172, App	922	33	19.3	340	2	US-08-636-883-4	Sequence 2, Appli
850	33	19.3	191	4	PCT-US95-10398-173	Sequence 173, App	923	33	19.3	359	1	US-07-662-005A-4	Sequence 4, Appli
851	33	19.3	191	4	PCT-US95-10398-174	Sequence 174, App	924	33	19.3	389	1	US-08-324-977-4	Sequence 4, Appli
852	33	19.3	191	4	PCT-US95-10398-175	Sequence 175, App	925	33	19.3	389	1	US-08-324-977-10	Sequence 10, Appl
853	33	19.3	191	4	PCT-US95-10398-176	Sequence 176, App	926	33	19.3	389	2	US-08-384-616-4	Sequence 4, Appli
854	33	19.3	191	4	PCT-US95-10398-177	Sequence 177, App	927	33	19.3	389	2	US-08-384-616-10	Sequence 10, Appl
855	33	19.3	191	4	PCT-US95-10398-178	Sequence 178, App	928	33	19.3	389	2	US-08-904-686A-4	Sequence 4, Appli
856	33	19.3	191	4	PCT-US95-10398-179	Sequence 179, App	929	33	19.3	389	2	US-08-904-686A-10	Sequence 10, Appl
857	33	19.3	191	4	PCT-US95-10398-180	Sequence 180, App	930	33	19.3	425	4	PCT-US92-06965A-27	Sequence 27, Appl
858	33	19.3	191	4	PCT-US95-10398-181	Sequence 181, App	931	33	19.3	428	1	US-08-103-445-2	Sequence 2, Appli
859	33	19.3	191	4	PCT-US95-10398-182	Sequence 182, App	932	33	19.3	428	1	US-08-461-690B-2	Sequence 2, Appli
860	33	19.3	191	4	PCT-US95-10398-183	Sequence 183, App	933	33	19.3	430	2	US-08-945-848-8	Sequence 8, Appli
861	33	19.3	191	4	PCT-US95-10398-184	Sequence 184, App	934	33	19.3	431	4	PCT-US91-00909-6	Sequence 6, Appli
862	33	19.3	191	4	PCT-US95-10398-185	Sequence 185, App	935	33	19.3	434	1	US-08-404-127-2	Sequence 2, Appli
863	33	19.3	191	4	PCT-US95-10398-186	Sequence 186, App	936	33	19.3	434	2	US-09-041-011-2	Sequence 2, Appli
864	33	19.3	191	4	PCT-US95-10398-187	Sequence 187, App	937	33	19.3	445	1	US-08-404-127-4	Sequence 4, Appli
865	33	19.3	191	4	PCT-US95-10398-188	Sequence 188, App	938	33	19.3	445	2	US-08-861-464-2	Sequence 2, Appli
866	33	19.3	191	4	PCT-US95-10398-189	Sequence 189, App	939	33	19.3	450	2	US-08-396-001-2	Sequence 2, Appli
867	33	19.3	191	4	PCT-US95-10398-190	Sequence 190, App	940	33	19.3	450	2	US-08-204-656B-2	Sequence 2, Appli
868	33	19.3	191	4	PCT-US95-10398-191	Sequence 191, App	941	33	19.3	468	1	US-08-204-656B-4	Sequence 4, Appli
869	33	19.3	191	4	PCT-US95-10398-192	Sequence 192, App	942	33	19.3	468	1	US-08-204-656B-6	Sequence 6, Appli
870	33	19.3	191	4	PCT-US95-10398-193	Sequence 193, App	943	33	19.3	468	1	US-08-204-656B-8	Sequence 8, Appli
871	33	19.3	191	4	PCT-US95-10398-194	Sequence 194, App	944	33	19.3	468	1	US-08-470-702-6	Sequence 6, Appli
872	33	19.3	191	4	PCT-US95-10398-195	Sequence 195, App	945	33	19.3	468	1	US-08-470-702-7	Sequence 7, Appli
873	33	19.3	191	4	PCT-US95-10398-196	Sequence 196, App	946	33	19.3	468	1	US-08-470-702-8	Sequence 8, Appli
874	33	19.3	191	4	PCT-US95-10398-197	Sequence 197, App	947	33	19.3	468	1	US-08-470-702-9	Sequence 9, Appli
875	33	19.3	191	4	PCT-US95-10398-198	Sequence 198, App	948	33	19.3	468	1	US-08-467-831-6	Sequence 6, Appli
876	33	19.3	191	4	PCT-US95-10398-199	Sequence 199, App	949	33	19.3	468	1	US-08-467-831-7	Sequence 7, Appli
877	33	19.3	192	1	US-08-086-428B-95	Sequence 95, Appl	950	33	19.3	468	1	US-08-467-831-8	Sequence 8, Appli
878	33	19.3	192	2	US-08-468-570-95	Sequence 95, Appl	951	33	19.3	468	1	US-08-467-831-9	Sequence 9, Appli
879	33	19.3	192	2	US-08-290-665A-95	Sequence 95, Appl	952	33	19.3	484	1	US-08-361-611-4	Sequence 4, Appli
880	33	19.3	192	4	PCT-US95-10398-200	Sequence 95, Appl	953	33	19.3	484	1	US-08-565-655-4	Sequence 4, Appli
881	33	19.3	201	2	US-08-531-525-13	Sequence 13, Appl	954	33	19.3	484	2	US-08-946-967-4	Sequence 4, Appli
882	33	19.3	201	2	US-08-916-901-8	Sequence 8, Appli	955	33	19.3	488	1	US-08-448-170-10	Sequence 10, Appl
883	33	19.3	201	2	US-08-718-270A-13	Sequence 13, Appl	956	33	19.3	488	3	US-08-484-661A-39	Sequence 39, Appl
884	33	19.3	202	2	US-08-531-525-14	Sequence 14, Appl	957	33	19.3	494	4	PCT-US96-09641-39	Sequence 39, Appl
885	33	19.3	202	2	US-08-718-270A-14	Sequence 14, Appl	958	33	19.3	505	1	US-08-222-616-20	Sequence 20, Appl
886	33	19.3	202	2	US-08-531-525-25	Sequence 25, Appl	959	33	19.3	505	4	PCT-US95-04228-20	Sequence 20, Appl
887	33	19.3	205	2	US-08-718-270A-25	Sequence 25, Appl	960	33	19.3	506	2	US-08-820-170A-19	Sequence 19, Appl
888	33	19.3	205	5	5185431-1	Patent No. 5185431	961	33	19.3	511	4	PCT-US95-05008-6	Sequence 6, Appli
889	33	19.3	206	5	US-08-824-873-4	Sequence 4, Appli	962	33	19.3	513	3	US-08-369-822C-28	Sequence 28, Appl
890	33	19.3	207	3	US-08-198-184-4	Sequence 4, Appli	963	33	19.3	532	1	US-08-188-228-44	Sequence 44, Appl
891	33	19.3	215	1	US-07-681-703B-12	Sequence 12, Appl	964	33	19.3	532	1	US-08-332-638-44	Sequence 44, Appl
892	33	19.3	215	1	US-08-531-525-10	Sequence 10, Appl	965	33	19.3	537	1	US-08-604-333-4	Sequence 4, Appli
893	33	19.3	215	2	US-08-718-270A-10	Sequence 10, Appl	966	33	19.3	537	3	PCT-US96-09641-37	Sequence 37, Appl
894	33	19.3	215	4	PCT-US91-02370-12	Sequence 12, Appl	967	33	19.3	571	4	US-08-484-661A-11	Sequence 11, Appl
895	33	19.3	215	4	PCT-US91-02370-12	Sequence 12, Appl	968	33	19.3	578	3	PCT-US96-09641-11	Sequence 11, Appl
896	33	19.3	215	4	PCT-US91-02370-12	Sequence 12, Appl	969	33	19.3	578	4	US-08-374-155A-14	Sequence 14, Appl
897	33	19.3	217	2	US-08-407-410B-12	Sequence 12, Appl	970	33	19.3	593	1	US-08-785-396-14	Sequence 14, Appl
898	33	19.3	217	2	US-08-407-410B-12	Sequence 12, Appl	971	33	19.3	593	2	US-08-799-138-4	Sequence 4, Appli
899	33	19.3	217	2	US-08-485-500-14	Sequence 14, Appl	972	33	19.3	610	3	US-08-484-661A-8	Sequence 8, Appli
900	33	19.3	217	2	US-08-485-500-14	Sequence 14, Appl	973	33	19.3	610	3	US-08-484-661A-16	Sequence 16, Appl
901	33	19.3	222	5	5185431-4	Patent No. 5185431	974	33	19.3	610	3	US-08-484-661A-19	Sequence 19, Appl
902	33	19.3	223	1	US-08-143-579A-4	Sequence 4, Appli	975	33	19.3	610	3	US-08-484-661A-23	Sequence 23, Appl
903	33	19.3	223	1	US-08-143-579A-4	Sequence 4, Appli	976	33	19.3	610	3		
904	33	19.3	223	1	US-08-143-579A-4	Sequence 4, Appli	977	33	19.3	610	3		

978 33 19.3 610 3 US-08-484-661A-29 Sequence 29, Appl
979 33 19.3 610 3 US-08-484-661A-33 Sequence 33, Appl
980 33 19.3 610 3 US-08-484-661A-35 Sequence 35, Appl
981 33 19.3 610 4 PCT-US96-09641-8 Sequence 8, Appl
982 33 19.3 610 4 PCT-US96-09641-16 Sequence 16, Appl
983 33 19.3 610 4 PCT-US96-09641-19 Sequence 19, Appl
984 33 19.3 610 4 PCT-US96-09641-23 Sequence 23, Appl
985 33 19.3 610 4 PCT-US96-09641-29 Sequence 29, Appl
986 33 19.3 610 4 PCT-US96-09641-33 Sequence 33, Appl
987 33 19.3 610 4 PCT-US96-09641-35 Sequence 35, Appl
988 33 19.3 610 4 PCT-US96-09641-54 Sequence 54, Appl
989 33 19.3 612 1 US-08-752-307B-11 Sequence 11, Appl
990 33 19.3 612 1 US-08-638-911A-35 Sequence 35, Appl
991 33 19.3 621 3 US-09-026-343-7 Sequence 7, Appl
992 33 19.3 668 1 US-08-205-018-2 Sequence 2, Appl
993 33 19.3 681 1 US-08-083-590A-18 Sequence 18, Appl
994 33 19.3 681 2 US-08-346-128-37 Sequence 37, Appl
995 33 19.3 735 2 US-08-313-185-48 Sequence 48, Appl
996 33 19.3 735 2 US-08-459-499-9 Sequence 9, Appl
997 33 19.3 735 2 US-08-459-499-12 Sequence 12, Appl
998 33 19.3 740 1 US-08-418-782-7 Sequence 7, Appl
999 33 19.3 740 1 US-08-418-782-21 Sequence 21, Appl
1000 33 19.3 740 1 US-08-228-662-7 Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-425-673-6
; Sequence 6, Application US/08425673
; Patent No. 5508261
; GENERAL INFORMATION:
; APPLICANT: Moyle, William R.
; APPLICANT: Campbell, Robert K.
; TITLE OF INVENTION: Analogs of Glycoprotein Hormones Having
; TITLE OF INVENTION: Altered Receptor Binding Specificity and Activity and
; TITLE OF INVENTION: Methods For Preparing and Using Same
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard R. Muccino
; STREET: P.O. Box 1267
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08551
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/425,673
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/17,151
; FILING DATE: 18-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Muccino, Richard R.
; REGISTRATION NUMBER: 32,538
; REFERENCE/DOCKET NUMBER: UMD 1.0-004
; TELEPHONE: (609) 466-3407
; TELEFAX: (609) 466-2760
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 112 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO

US-08-425-673-6
Query Match 95.3%; Score 163; DB 1; Length 112;
Best Local Similarity 96.7%; Pred. No. 2.7e-16;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 vctyrdfiyrtveipgclphvapyfsypva 30
Db 51 VCTYRDFIYRTVEIPQCPLHVPYFSPVA 80
RESULT 2
US-08-239-256-3
; Sequence 3, Application US/08239256
; Patent No. 5585345
; GENERAL INFORMATION:
; APPLICANT: BOIME, IRVING
; APPLICANT: MATZUK, MARTIN M.
; APPLICANT: KEENE, JEFFREY L.
; TITLE OF INVENTION: CTP EXTENDED FORM OF LH
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/239,256
; FILING DATE: 06-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 29500-20030.12
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 75 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; APPLICATION NUMBER: US-08-239-256-3

Query Match 53.2%; Score 91; DB 1; Length 75;
Best Local Similarity 44.8%; Pred. No. 2.7e-06;
Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;
Qy 2 ctyrdfiyrtveipgclphvapyfsypva 30
Db 16 CTFKELVYETVRVPGCAHADSLSYTPVA 44
RESULT 3
517193-5
; Patent No. 5177193
; APPLICANT: BOIME, IRVING; MATZUK, MARTIN M.
; TITLE OF INVENTION: MODIFIED FORMS OF REPRODUCTIVE HORMONES
; NUMBER OF SEQUENCES: 26
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/532,254
; FILING DATE: 01-JUN-1990
; PRIOR APPLICATION DATA:

Query Match 53.2%; Score 91; DB 5; Length 131;
Best Local Similarity 44.8%; Pred. No. 5e-06; 9; Mismatches 0; Gaps 0;
Matches 13; Conservative 7;

Qy 2 ctgyrdflvrtveipgpcplhvapfsypva 30
|||: :||: ||| :|||
Db 70 CTRELVTYTVRVPCAHADSLTYTPVA 98

RESULT 7
US-08-425-673-5
; Sequence 5, Application US/08425673
; Patent No. 5508261
; GENERAL INFORMATION:
; APPLICANT: Moyle, William R.
; APPLICANT: Campbell, Robert K.
; TITLE OF INVENTION: Analogs of Glycoprotein Hormones Having
; TITLE OF INVENTION: Altered Receptor Binding Specificity and Activity and
; TITLE OF INVENTION: Methods For Preparing and Using Same
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard R. Muccino
; STREET: P.O. Box 1267
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08551
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/425,673
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/717,151
; FILING DATE: 18-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Muccino, Richard R.
; REGISTRATION NUMBER: 32,538
; REFERENCE/DOCKET NUMBER: UMD 1.0-004
; TELEPHONE: (609) 466-3407
; TELEFAX: (609) 466-2760
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 149 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-425-673-5

Query Match 52.6%; Score 90; DB 1; Length 149;
Best Local Similarity 50.0%; Pred. No. 8e-06; 9; Indels 0; Gaps 0;
Matches 15; Conservative 6; Mismatches 9;

Qy 1 vctyrdflvrtveipgpcplhvapfsypva 30
|||: :||: ||| :|||
Db 56 VCTRELRFASIRLPGCPGVPWVSPVA 85

RESULT 8
US-08-327-362-2
; Sequence 2, Application US/08327362
; Patent No. 5811249
; GENERAL INFORMATION:
; APPLICANT: William D. Odell, Jeanine T. Griffin, Sanjeev

Query Match 50.3%; Score 86; DB 2; Length 88;
Best Local Similarity 50.0%; Pred. No. 1.6e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 1 vctyrdflvrtveipgpcplhvapfsypva 30
|||: :||: ||| :|||
Db 5 VCNVRDVRFSIRLPGCPGVPWVSYAVA 34

RESULT 9
US-08-425-673-7
; Sequence 7, Application US/08425673
; Patent No. 5508261
; GENERAL INFORMATION:
; APPLICANT: Moyle, William R.
; APPLICANT: Campbell, Robert K.
; TITLE OF INVENTION: Analogs of Glycoprotein Hormones Having
; TITLE OF INVENTION: Altered Receptor Binding Specificity and Activity and
; TITLE OF INVENTION: Methods For Preparing and Using Same
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard R. Muccino
; STREET: P.O. Box 1267
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08551
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
US-08-425-673-5

Query Match 50.3%; Score 86; DB 2; Length 88;
Best Local Similarity 50.0%; Pred. No. 1.6e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 1 vctyrdflvrtveipgpcplhvapfsypva 30
|||: :||: ||| :|||
Db 5 VCNVRDVRFSIRLPGCPGVPWVSYAVA 34

RESULT 9
US-08-425-673-7
; Sequence 7, Application US/08425673
; Patent No. 5508261
; GENERAL INFORMATION:
; APPLICANT: Moyle, William R.
; APPLICANT: Campbell, Robert K.
; TITLE OF INVENTION: Analogs of Glycoprotein Hormones Having
; TITLE OF INVENTION: Altered Receptor Binding Specificity and Activity and
; TITLE OF INVENTION: Methods For Preparing and Using Same
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard R. Muccino
; STREET: P.O. Box 1267
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08551
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
US-08-327-362-2

APPLICANT: Grover, Omar Caticha, Douglas T. Carrell,
APPLICANT: Marion L. Woods
TITLE OF INVENTION: Control of Infectious Microorganisms
TITLE OF INVENTION: by Modulation of Choriionic
TITLE OF INVENTION: Gonadotropin-Related Protein
TITLE OF INVENTION: Activity
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Thorpe, No. 5811249th & Western
STREET: 9035 South 700 East, Suite 200
CITY: Sandy
STATE: Utah
COUNTRY: USA
ZIP: 84070
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb storage
COMPUTER: AST Advantage NB-SX20
OPERATING SYSTEM: DOS 6.1
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/327,362
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Alan J. Howarth
REGISTRATION NUMBER: 36,553
REFERENCE/DOCKET NUMBER: T1893
TELECOMMUNICATION INFORMATION:
TELEPHONE: (801)566-6633
TELEFAX: (801)566-0750
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 88 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-327-362-2

Query Match 50.3%; Score 86; DB 2; Length 88;
Best Local Similarity 50.0%; Pred. No. 1.6e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 1 vctyrdflvrtveipgpcplhvapfsypva 30
|||: :||: ||| :|||
Db 5 VCNVRDVRFSIRLPGCPGVPWVSYAVA 34

RESULT 9
US-08-425-673-7
; Sequence 7, Application US/08425673
; Patent No. 5508261
; GENERAL INFORMATION:
; APPLICANT: Moyle, William R.
; APPLICANT: Campbell, Robert K.
; TITLE OF INVENTION: Analogs of Glycoprotein Hormones Having
; TITLE OF INVENTION: Altered Receptor Binding Specificity and Activity and
; TITLE OF INVENTION: Methods For Preparing and Using Same
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard R. Muccino
; STREET: P.O. Box 1267
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08551
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
US-08-327-362-2

Query Match 50.3%; Score 86; DB 2; Length 88;
Best Local Similarity 50.0%; Pred. No. 1.6e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 1 vctyrdflvrtveipgpcplhvapfsypva 30
|||: :||: ||| :|||
Db 5 VCNVRDVRFSIRLPGCPGVPWVSYAVA 34

RESULT 9
US-08-425-673-7
; Sequence 7, Application US/08425673
; Patent No. 5508261
; GENERAL INFORMATION:
; APPLICANT: Moyle, William R.
; APPLICANT: Campbell, Robert K.
; TITLE OF INVENTION: Analogs of Glycoprotein Hormones Having
; TITLE OF INVENTION: Altered Receptor Binding Specificity and Activity and
; TITLE OF INVENTION: Methods For Preparing and Using Same
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard R. Muccino
; STREET: P.O. Box 1267
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08551
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
US-08-327-362-2

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/425,673
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/717,151
FILING DATE: 18-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Muccino, Richard R.
REGISTRATION NUMBER: 32,538
REFERENCE/DOCKET NUMBER: UMD 1.0-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 466-3407
TELEFAX: (609) 466-2760
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 114 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-425-673-7

Query Match 50.3%; Score 86; DB 1; Length 114;
Best Local Similarity 50.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpvhvafsyypva 30
|| ||| : : : |||| | | || ||
Db 56 VCNVYRDVRFESIRLPGCPGVNPDWSYAVA 85

RESULT 10
US-08-425-673-8
Sequence 8, Application US/08425673
Patent No. 5508261
GENERAL INFORMATION:
APPLICANT: Moyle, William R. K.
APPLICANT: Campbell, Robert K.
TITLE OF INVENTION: Analogs of Glycoprotein Hormones Having
TITLE OF INVENTION: Altered Receptor Binding Specificity and Activity and
TITLE OF INVENTION: Methods For Preparing and Using Same
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richard R. Muccino
STREET: P.O. Box 1267
CITY: Princeton
STATE: New Jersey
COUNTRY: USA
ZIP: 08551
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/425,673
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/717,151
FILING DATE: 18-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Muccino, Richard R.
REGISTRATION NUMBER: 32,538
REFERENCE/DOCKET NUMBER: UMD 1.0-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 466-3407

TELEFAX: (609) 466-2760
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 114 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-425-673-8

Query Match 50.3%; Score 86; DB 1; Length 114;
Best Local Similarity 50.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpvhvafsyypva 30
|| ||| : : : |||| | | || ||
Db 56 VCNVYRDVRFESIRLPGCPGVNPDWSYAVA 85

RESULT 11
US-08-425-673-9
Sequence 9, Application US/08425673
Patent No. 5508261
GENERAL INFORMATION:
APPLICANT: Moyle, William R. K.
APPLICANT: Campbell, Robert K.
TITLE OF INVENTION: Analogs of Glycoprotein Hormones Having
TITLE OF INVENTION: Altered Receptor Binding Specificity and Activity and
TITLE OF INVENTION: Methods For Preparing and Using Same
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richard R. Muccino
STREET: P.O. Box 1267
CITY: Princeton
STATE: New Jersey
COUNTRY: USA
ZIP: 08551
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/425,673
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/717,151
FILING DATE: 18-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Muccino, Richard R.
REGISTRATION NUMBER: 32,538
REFERENCE/DOCKET NUMBER: UMD 1.0-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 466-3407
TELEFAX: (609) 466-2760
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 114 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-425-673-9

Query Match 50.3%; Score 86; DB 1; Length 114;
Best Local Similarity 50.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;


```
Qy 1 vctyrdfiyrtveipgcpplhvapyfsypva 30
|| ||| : : : |||| | | || ||
Db 56 VCNVDRVRFESIRLPGCPGVNPVSYAVA 85

RESULT 12
US-08-425-673-11
; Sequence 11, Application US/08425673
; Patent No. 5508261
; GENERAL INFORMATION:
; APPLICANT: Moyle, William R.
; TITLE OF INVENTION: Analogs of Glycoprotein Hormones Having
; TITLE OF INVENTION: Altered Receptor Binding Specificity and Activity and
; TITLE OF INVENTION: Methods For Preparing and Using Same
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard R. Muccino
; STREET: P.O. Box 1267
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08551
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/425,673
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/717,151
; FILING DATE: 18-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Muccino, Richard R.
; REGISTRATION NUMBER: 32,538
; REFERENCE/DOCKET NUMBER: UMD 1.0-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 466-3407
; TELEFAX: (609) 466-2760
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-425-673-11

Query Match 50.3%; Score 86; DB 1; Length 117;
Best Local Similarity 50.0%; Pred. No. 2.3e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcpplhvapyfsypva 30
|| ||| : : : |||| | | || ||
Db 56 VCNVDRVRFESIRLPGCPGVNPVSYAVA 85

RESULT 13
US-08-425-673-12
; Sequence 12, Application US/08425673
; Patent No. 5508261
; GENERAL INFORMATION:
; APPLICANT: Moyle, William R.
; TITLE OF INVENTION: Analogs of Glycoprotein Hormones Having
; TITLE OF INVENTION: Altered Receptor Binding Specificity and Activity and
; TITLE OF INVENTION: Methods For Preparing and Using Same
; NUMBER OF SEQUENCES: 12
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard R. Muccino
; STREET: P.O. Box 1267
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08551
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/425,673
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/717,151
; FILING DATE: 18-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Muccino, Richard R.
; REGISTRATION NUMBER: 32,538
; REFERENCE/DOCKET NUMBER: UMD 1.0-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 466-3407
; TELEFAX: (609) 466-2760
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-425-673-12

Query Match 50.3%; Score 86; DB 1; Length 117;
Best Local Similarity 50.0%; Pred. No. 2.3e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcpplhvapyfsypva 30
|| ||| : : : |||| | | || ||
Db 56 VCNVDRVRFESIRLPGCPGVNPVSYAVA 85

RESULT 14
US-08-425-673-1
; Sequence 1, Application US/08425673
; Patent No. 5508261
; GENERAL INFORMATION:
; APPLICANT: Moyle, William R.
; APPLICANT: Campbell, Robert K.
; TITLE OF INVENTION: Analogs of Glycoprotein Hormones Having
; TITLE OF INVENTION: Altered Receptor Binding Specificity and Activity and
; TITLE OF INVENTION: Methods For Preparing and Using Same
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard R. Muccino
; STREET: P.O. Box 1267
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08551
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/425,673
; FILING DATE:
; CLASSIFICATION: 514
```


;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/717,151
;; FILING DATE: 18-JUN-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Muccino, Richard R.
;; REGISTRATION NUMBER: 32,538
;; REFERENCE/DOCKET NUMBER: UMD 1.0-004
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (609) 466-3407
;; TELEFAX: (609) 466-2760
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 145 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-08-425-673-1

Query Match 50.3%; Score 86; DB 1; Length 145;
Best Local Similarity 50.0%; Pred. No. 2.9e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpvhvafsyypva 30
|| ||| : : : |||| | | || ||
Db 56 VCNVDRVRFESIRLPGCPRGVNPVVSAYA 85

RESULT 15
US-08-425-673-2
;; Sequence 2, Application US/08425673
;; Patent No. 5508261
;; GENERAL INFORMATION:
;; APPLICANT: Moyle, William R.
;; APPLICANT: Campbell, Robert K.
;; TITLE OF INVENTION: Analogs of Glycoprotein Hormones Having
;; TITLE OF INVENTION: Altered Receptor Binding Specificity and Activity and
;; TITLE OF INVENTION: Methods For Preparing and Using Same
;; NUMBER OF SEQUENCES: 12
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Richard R. Muccino
;; STREET: P.O. Box 1267
;; CITY: Princeton
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 08551
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/425,673
;; FILING DATE:
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/717,151
;; FILING DATE: 18-JUN-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Muccino, Richard R.
;; REGISTRATION NUMBER: 32,538
;; REFERENCE/DOCKET NUMBER: UMD 1.0-004
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (609) 466-3407
;; TELEFAX: (609) 466-2760
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 145 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide

;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-08-425-673-2

Query Match 50.3%; Score 86; DB 1; Length 145;
Best Local Similarity 50.0%; Pred. No. 2.9e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpvhvafsyypva 30
|| ||| : : : |||| | | || ||
Db 56 VCNVDRVRFESIRLPGCPRGVNPVVSAYA 85

RESULT 16
US-08-425-673-10
;; Sequence 10, Application US/08425673
;; Patent No. 5508261
;; GENERAL INFORMATION:
;; APPLICANT: Moyle, William R.
;; APPLICANT: Campbell, Robert K.
;; TITLE OF INVENTION: Analogs of Glycoprotein Hormones Having
;; TITLE OF INVENTION: Altered Receptor Binding Specificity and Activity and
;; TITLE OF INVENTION: Methods For Preparing and Using Same
;; NUMBER OF SEQUENCES: 12
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Richard R. Muccino
;; STREET: P.O. Box 1267
;; CITY: Princeton
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 08551
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/425,673
;; FILING DATE:
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/717,151
;; FILING DATE: 18-JUN-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Muccino, Richard R.
;; REGISTRATION NUMBER: 32,538
;; REFERENCE/DOCKET NUMBER: UMD 1.0-004
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (609) 466-3407
;; TELEFAX: (609) 466-2760
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 145 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-08-425-673-10

Query Match 50.3%; Score 86; DB 1; Length 145;
Best Local Similarity 50.0%; Pred. No. 2.9e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpvhvafsyypva 30
|| ||| : : : |||| | | || ||
Db 56 VCNVDRVRFESIRLPGCPRGVNPVVSAYA 85

RESULT 17
US-08-298-189B-1

; Sequence 1, Application US/08298189B
; Patent No. 5674727
; GENERAL INFORMATION:
; APPLICANT: Dr. Laurence A. Cole and Dr. Andrew Kardana
; TITLE OF INVENTION: No. 5674727el Methods for Detecting Reproductive
; TITLE OF INVENTION: Cancers or Tumors and Assay Products
; TITLE OF INVENTION: Therefor
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Laurence A. Cole, c/o Dept. of Obstetrics and
; STREET: 333 Cedar Street
; CITY: New Haven
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06510
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Microsoft Word 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/298,189B
; FILING DATE: 08/31/94
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brian D. Voyce
; REGISTRATION NUMBER: 28,917
; REFERENCE/DOCKET NUMBER: DS11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-638-3939 or 803-272-1471
; TELEFAX: 919-638-3939 or 803-272-1471
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. 5674727 applicable
; TOPOLOGY: Unknown
; MOLECULE TYPE: subunit of hormone, specifically the
; MOLECULE TYPE: beta\subunit of hCG
; ORIGINAL SOURCE: human urine
; FEATURE:
; NAME/KEY: beta subunit of hCG that is nicked by GBNE
; LOCATION: hCG
; IDENTIFICATION METHOD: N-terminal sequence analysis
; PUBLICATION INFORMATION:
; AUTHORS: Keutmann et alia
; TITLE: "A Receptor-binding Region in Human
; JOURNAL: Proc Nat'l Acad Sci USA
; VOLUME: 84
; ISSUE: No. 5674727 applicable
; PAGES: 2038-2042
; DATE: 1987
; US-08-298-189B-1

Query Match 50.3%; Score 86; DB 1; Length 145;
Best Local Similarity 50.0%; Pred. No. 2.9e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 1 vctyrdfyrtveipgcpvhvapyfsypva 30
||| | | : : : ||| | | | | |
Db 56 VCNTRDVRFSIRLPGCPGVNPNVSVAYA 85

RESULT 18
US-08-475-213-10
; Sequence 10, Application US/08475213
; Patent No. 5785674
; GENERAL INFORMATION:
; APPLICANT: Geysen, Hendrik M.
; TITLE OF INVENTION: Method for the use and Synthesis of
; TITLE OF INVENTION: Peptides
; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,213
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/752,437
; FILING DATE: 06-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO pct/au90/00062
; FILING DATE: 16-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PJ2788/89
; FILING DATE: 17-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31259
; REFERENCE/DOCKET NUMBER: 0240.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-475-213-10

Query Match 50.3%; Score 86; DB 1; Length 145;
Best Local Similarity 50.0%; Pred. No. 2.9e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 1 vctyrdfyrtveipgcpvhvapyfsypva 30
||| | | : : : ||| | | | | |
Db 56 VCNTRDVRFSIRLPGCPGVNPNVSVAYA 85

RESULT 19
US-08-395-238-2
; Sequence 2, Application US/08395238
; Patent No. 5864488
; GENERAL INFORMATION:
; APPLICANT: ISSACS, Neil William
; APPLICANT: LAPHORN, Adrian Jonathan
; APPLICANT: HARRIS, Deborah Claire
; TITLE OF INVENTION: THREE DIMENSIONAL HORMONE STRUCTURE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AKZO NOBEL PATENT DEPARTMENT
; STREET: 1300 PICCARD DRIVE, SUITE 206
; CITY: ROCKVILLE
; STATE: MARYLAND
; COUNTRY: UNITED STATES
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,238
; FILING DATE: 24-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9403600.1
; FILING DATE: 24-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM M. BLACKSTONE
; REGISTRATION NUMBER: 29,722
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; STRAIN: BETA-SUBUNIT HUMAN CHORIONIC GONADOTROPIN
; US-08-395-238-2

Query Match 50.3%; Score 86; DB 2; Length 145;
Best Local Similarity 50.0%; Pred. No. 2.9e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 vctyrdfltyrtveipgcpplhvapfysypva 30
|| || | : : : || || | : : || || |
Db 56 VCNVYRDVRFESIRLPGCPGVNPFVSYAVA 85

RESULT 20
US-08-709-924-23
; Sequence 23, Application US/08709924
; Patent No. 5968513
; GENERAL INFORMATION:
; APPLICANT: Gallo, Robert C.
; APPLICANT: Bryant, Joseph
; APPLICANT: Lunardi-Iskandar, Yanto
; TITLE OF INVENTION: METHODS OF PROMOTING HEMATOPOIESIS
; TITLE OF INVENTION: USING DERIVATIVES OF HUMAN CHORIONIC GONADOTROPIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,924
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 8769-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
; US-08-709-924-23

Query Match 48.5%; Score 83; DB 2; Length 98;
Best Local Similarity 46.7%; Pred. No. 4.9e-05;
Matches 14; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 vctyrdfltyrtveipgcpplhvapfysypva 30
|| || | : : : || || | : : || || |
Db 9 VCNVYRDVRFESIRLPGCPGLNPFVSYAVA 38

RESULT 21
US-08-709-925-23
; Sequence 23, Application US/08709925
; Patent No. 5997871
; GENERAL INFORMATION:
; APPLICANT: Gallo, Robert C.
; APPLICANT: Bryant, Joseph
; APPLICANT: Lunardi-Iskandar, Yanto
; TITLE OF INVENTION: TREATMENT AND PREVENTION OF CANCER BY
; TITLE OF INVENTION: ADMINISTRATION OF DERIVATIVES OF HUMAN CHORIONIC GONADOTROPIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,925
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 512
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 8769-017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-709-925-23

Query Match 48.5%; Score 83; DB 2; Length 98;
Best Local Similarity 46.7%; Pred. No. 4.9e-05;
Matches 14; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 vctyrdfltyrtveipgcpplhvapfysypva 30
|| || | : : : || || | : : || || |
Db 9 VCNVYRDVRFESIRLPGCPGLNPFVSYAVA 38

RESULT 22
US-08-709-924-2
; Sequence 2, Application US/08709924
; Patent No. 5968513
; GENERAL INFORMATION:
; APPLICANT: Gallo, Robert C.
; APPLICANT: Bryant, Joseph


```
; APPLICANT: Lunardi-Iskandar, Yanto
; TITLE OF INVENTION: METHODS OF PROMOTING HEMATOPOIESIS
; REFERENCE/DOCKET NUMBER: 8769-018
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,924
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 8769-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-709-924-2
;
; Query Match 48.5%; Score 83; DB 2; Length 165;
; Best Local Similarity 46.7%; Pred. No. 8.8e-05;
; Matches 14; Conservative 5; Mismatches 11; Indels 0; Gaps 0;
;
QY 1 vctyrdfiyrtveipgpcplhvapysypva 30
   || ||| : : : |||| : | || ||
Db 76 VCNYRDVRFESIRLPGCPRLNPVYSYAVA 105

RESULT 23
US-08-709-925-2
; Sequence 2, Application US/08709925
; Patent No. 5997871
; GENERAL INFORMATION:
; APPLICANT: Gallo, Robert C.
; APPLICANT: Bryant, Joseph
; APPLICANT: Lunardi-Iskandar, Yanto
; TITLE OF INVENTION: TREATMENT AND PREVENTION OF CANCER BY
; ADMINISTRATION OF DERIVATIVES OF HUMAN CHORIONIC GONADOTROPIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,925
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 514
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 8769-017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-709-925-2
;
; Query Match 48.5%; Score 83; DB 2; Length 165;
; Best Local Similarity 46.7%; Pred. No. 8.8e-05;
; Matches 14; Conservative 5; Mismatches 11; Indels 0; Gaps 0;
;
QY 1 vctyrdfiyrtveipgpcplhvapysypva 30
   || ||| : : : |||| : | || ||
Db 76 VCNYRDVRFESIRLPGCPRLNPVYSYAVA 105

RESULT 24
US-08-709-924-24
; Sequence 24, Application US/08709924
; Patent No. 5968513
; GENERAL INFORMATION:
; APPLICANT: Gallo, Robert C.
; APPLICANT: Bryant, Joseph
; APPLICANT: Lunardi-Iskandar, Yanto
; TITLE OF INVENTION: METHODS OF PROMOTING HEMATOPOIESIS
; USING DERIVATIVES OF HUMAN CHORIONIC GONADOTROPIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,924
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 8769-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 88 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-709-924-24
;
; Query Match 40.9%; Score 70; DB 2; Length 88;
; Best Local Similarity 44.4%; Pred. No. 0.003;
; Matches 12; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
```


QY 4 yrdfiyrtveipgcpvhvapyfsypva 30
||| : : : |||| : | || ||
Db 2 YRDVRESIRLPGCPRLNPVVSIAVA 28

RESULT 25

US-08-709-925-24
; Sequence 24, Application US/08709925
; Patent No. 5997871
; GENERAL INFORMATION:
; APPLICANT: Gallo, Robert C.
; APPLICANT: Bryant, Joseph
; APPLICANT: Lunardi-Iskandar, Yanto
; TITLE OF INVENTION: TREATMENT AND PREVENTION OF CANCER BY
; ADMINISTRATION OF DERIVATIVES OF HUMAN CHORIONIC GONADOTROPIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,925
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 512
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 8769-017
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 88 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-709-925-24

Query Match 40.9%; Score 70; DB 2; Length 88;
Best Local Similarity 44.4%; Pred. No. 0.003;
Matches 12; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 4 yrdfiyrtveipgcpvhvapyfsypva 30
||| : : : |||| : | || ||
Db 2 YRDVRESIRLPGCPRLNPVVSIAVA 28

RESULT 26

US-08-036-555B-20
; Sequence 20, Application US/08036555B
; Patent No. 5530109
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew; Stroobant, Paul;
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
; APPLICANT: Chen, Malo Su; Hiles, Ian
; TITLE OF INVENTION: Glial Mitogenic Factors, Their
; PREPARATION AND USE
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022

; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/036,555B
; FILING DATE: 24-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APRIL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LUD 5250.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-036-555B-20

Query Match 36.3%; Score 62; DB 1; Length 26;
Best Local Similarity 50.0%; Pred. No. 0.011;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 9 yrtveipgcpvhvapyfsypva 30
: | : |||| | | : ||||
Db 4 FASVRLPGCPGVDPMVSPVA 25

RESULT 27

US-08-469-569-20
; Sequence 20, Application US/08469569
; Patent No. 5606032
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew; Stroobant, Paul;
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
; APPLICANT: Chen, Malo Su; Hiles, Ian
; TITLE OF INVENTION: Glial Mitogenic Factors, Their
; PREPARATION AND USE
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage


```
;
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,569
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 24-MAR-1993
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APRIL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LUD 5250.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-469-569-20

Query Match 36.3%; Score 62; DB 1; Length 26;
Best Local Similarity 50.0%; Pred. No. 0.011;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 9 yrtveipgpcplhvapyfsypva 30
Db 4 FASVRLPGCPGVDPMVSFPVA 25
```

```
RESULT 28
US-08-249-322A-20
; Sequence 20, Application US/08249322A
; Patent No. 5716930
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew; Stroobant, Paul;
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;
; APPLICANT: Chen, Maio Su; Hiles, Ian
; TITLE OF INVENTION: Glial Mitogenic Factors, Their
; PREPARATION AND USE
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
```

```
;
; APPLICATION NUMBER: US/08/249,322A
; FILING DATE: 26-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 24-MAR-1993
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APRIL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LUD 250.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-249-322A-20

Query Match 36.3%; Score 62; DB 1; Length 26;
Best Local Similarity 50.0%; Pred. No. 0.011;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 9 yrtveipgpcplhvapyfsypva 30
Db 4 FASVRLPGCPGVDPMVSFPVA 25

RESULT 29
US-08-469-526A-20
; Sequence 20, Application US/08469526A
; Patent No. 5792849
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew
; APPLICANT: Stroobant, Paul
; APPLICANT: Minghetti, Luisa
; APPLICANT: Waterfield, Michael
; APPLICANT: Marchionni, Mark
; APPLICANT: Chen, Maio Su
; APPLICANT: Hiles, Ian
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; PREPARATION AND USE
; NUMBER OF SEQUENCES: 187
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
```



```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,526A
; FILING DATE: 06 June 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 24-MAR-1993
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 03-JUN-1992
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bleker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 04585/00200A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-469-526A-20

Query Match          36.3%; Score 62; 1
Best Local Similarity 50.0%; Pred. No. 0
Matches 11; Conservative 4; Mismatch

OY   9 yrtveipgclphvapfysypva 30
      :|:||||||| |:|||
Db    4 FASVRLPGCPGVDPMPVPVA 25

RESULT 30
US-08-734-591A-20
; Sequence 20, Application US/08734591A
; Patent No. 5854220
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew
; APPLICANT: Stroobant, Paul
; APPLICANT: Minghettil, Luisa
; APPLICANT: Waterfield, Michael
; APPLICANT: Hiles, Ian
; APPLICANT: Marchionni, Mark
; APPLICANT: Chen, Mario
; TITLE OF INVENTION: GLIAL MITOGENIC FACTOR
; TITLE OF INVENTION: PREPARATION AND USE
; NUMBER OF SEQUENCES: 187
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible Pentium
; OPERATING SYSTEM: Windows95
; SOFTWARE: WordPerfect (Version 7.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,591A
; FILING DATE: 22-OCT-1996
; CLASSIFICATION: 536

```

```

; PRIOR APPLICATION DATA: 08/470,335
; APPLICATION NUMBER: 08/470,335
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 03-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 91 07566.3
; FILING DATE: 10-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bieker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 04585/00200P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 428-0200
; TELEFAX: (617) 428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-734-591A-20

Query Match 36.3%; Score 62; DB 2; Length 26;
Best Local Similarity 50.0%; Pred. No. 0.011;
Matches 11; Conservative 4; Mismatches 7; Indels

Qy 9 yrtveipgclhvpayfsypva 30
   : : : : : : : : : : : : : : : :
Db 4 FASVRLPGCPGVDPMVSPFVA 25

RESULT 31
US-08-469-660-20
; Sequence 20, Application US/08469660
; Patent No. 5876973
; GENERAL INFORMATION:
; APPLICANT: Gwynne, David I.; Marchionni, Mark;
; APPLICANT: McBurney, Robert N.
; TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,
; TITLE OF INVENTION: THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; ZIP: 0211-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,660
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```



```

; APPLICATION NUMBER: 08/011,396
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/984,085
; FILING DATE: 01-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/951,747
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/927,337
; FILING DATE: 10-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 04585/017004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: 200154
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-469-660-20

Query Match 36.3%; Score 62; 1
Best Local Similarity 50.0%; Pred. No. 0
Matches 11; Conservative 4; Mismatch

Qy 9 yrtveipgpcplhvapyfsypva 30
: : : : : : : : : : : : : :
Db 4 FASVRLPGCPGVDPMVSEFVA 25

RESULT 32
PCT-US94-05083C-20
; Sequence 20, Application PC/TUS9405083C
; GENERAL INFORMATION:
; APPLICANT: Robert Sklar, Mark Marchionn
; APPLICANT: David I. Gwynne
; TITLE OF INVENTION: METHODS FOR ALTERIN
; TITLE OF INVENTION: MUSCLE CONDITION
; NUMBER OF SEQUENCES: 185
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360
; MEDIUM TYPE: kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05083C
; FILING DATE: 06-MAY-94
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/209,204
; FILING DATE: 08-MAR-94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059,022
; FILING DATE: 06-MAY-93
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 04585/028W01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070

```


TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US95-06846A-20

Query Match 36.3%; Score 62; DB 4; Length 26;
Best Local Similarity 50.0%; Pred. No. 0.011;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 9 yrtveipgcpvhvpyfsypva 30
:|:||||| | | | | |
Db 4 FASVRLPGCPGVDPMVFPVA 25

RESULT 34
US-08-599-171A-27
; Sequence 27, Application US/08599171A
; Patent No. 5814473
; GENERAL INFORMATION:
; APPLICANT: WARREN, Patrick V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,171A
; FILING DATE: Concurrently
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HERRON, CHARLES J.
; REGISTRATION NUMBER: 28,019
; REFERENCE/DOCKET NUMBER: 331400-38
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 AMINO ACIDS
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-599-171A-27

Query Match 27.5%; Score 47; DB 2; Length 453;
Best Local Similarity 31.8%; Pred. No. 33;
Matches 7; Conservative 8; Mismatches 1; Indels 1; Gaps 1;

QY 3 tyrdfiyrtveipgcpvhvpy 24
||:|:||||:| | | | |
Db 170 TYKDLLEFKTIKLP-----SPY 185

RESULT 35
US-08-646-590B-27
; Sequence 27, Application US/08646590B
; Patent No. 5962283
; GENERAL INFORMATION:
; APPLICANT: Warren, Patrick V.
; APPLICANT: Swanson, Ronald V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646.590B
; FILING DATE: 08-May-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/599,171
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01094
; FILING DATE: 21-January-1997
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/017001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-646-590B-27

Query Match 27.5%; Score 47; DB 2; Length 453;
Best Local Similarity 31.8%; Pred. No. 33;
Matches 7; Conservative 8; Mismatches 1; Indels 6; Gaps 1;

QY 3 tyrdfiyrtveipgcpvhvpy 24
||:|:||||:| | | | |
Db 170 TYKDLLEFKTIKLP-----SPY 185

RESULT 36
US-09-069-226-27
; Sequence 27, Application US/09069226
; Patent No. 6013509
; GENERAL INFORMATION:
; APPLICANT: WARREN, Patrick V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA

ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/069,226
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/599,171
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HERRON, CHARLES J.
REGISTRATION NUMBER: 28,019
REFERENCE/DOCKET NUMBER: 331400-38
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 453 AMINO ACIDS
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-09-069-226-27

Query Match 27.5%; Score 47; DB 3; Length 453;
Best Local Similarity 31.8%; Pred. No. 33;
Matches 7; Conservative 8; Mismatches 1; Indels 6; Gaps 1;

QY 3 tyrdflrtveipgcplhvapy 24
||| :|:|:| :||
DB 170 TYKDLLEFKTKLP-----SPY 185

RESULT 37
US-07-593-657-7
Sequence 7, Application US/07593657
Patent No. 5266317
GENERAL INFORMATION:
APPLICANT: Tomalski, Michael D.
APPLICANT: Miller, Lois K.
TITLE OF INVENTION: INSECT-SPECIFIC PARALYTIC NEUROTOXIN
TITLE OF INVENTION: GENES FOR USE IN BIOLOGICAL INSECT CONTROL: METHODS AND
TITLE OF INVENTION: COMPOSITIONS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: Greenlee and Winner, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/593,657
FILING DATE: 19901004
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Feiber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 14-90
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/499-8080
TELEFAX: 303/499-8089
INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-593-657-7

Query Match 27.2%; Score 46.5; DB 1; Length 289;
Best Local Similarity 28.2%; Pred. No. 24;
Matches 11; Conservative 5; Mismatches 10; Indels 13; Gaps 1;

QY 2 ctyrdfi-----yrtveipgcplhvapyfsy 27
| | | | | : | | | : |
DB 117 CVERDFIEFGVSIKIGISHRGDDPACPLYENTWINY 155

RESULT 38
US-08-482-577B-27
Sequence 27, Application US/08482577B
Patent No. 5807713
GENERAL INFORMATION:
APPLICANT: HOTTEN, GERTRUD
APPLICANT: NEIDHARDT, HELGE
APPLICANT: BECHTOLD, ROLF
APPLICANT: POHL, JENS
TITLE OF INVENTION: DNA SEQUENCES ENCODING NOVEL
TITLE OF INVENTION: GROWTH/DIFFERENTIATION FACTORS
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESS: NIKAI, DAVID, MARCELSTEIN, MURRAY, AND ORAM
STREET: 655 FIFTEENTH STREET, N.W., G STREET LOBBY,
CITY: SUITE 330
STATE: WASHINGTON
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,577B
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KLESNER, SHARON
REGISTRATION NUMBER: 36,335
REFERENCE/DOCKET NUMBER: P564-5010
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/638-5000
TELEFAX: 202/638-4810
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-482-577B-27

Query Match 26.9%; Score 46; DB 1; Length 105;
Best Local Similarity 57.1%; Pred. No. 9.2;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 16 gcplhvapyfsypv 29
| | | | | : | | | : |
DB 33 GCGLHPPNLSLPV 46

RESULT 39


```
US-08-478-097A-22
; Sequence 22, Application US/08478097A
; Patent No. 6040431
; GENERAL INFORMATION:
; APPLICANT: KECK, PETER
; APPLICANT: SMART, JOHN
; TITLE OF INVENTION: SINGLE-CHAIN ANALOGS OF TGF-B
; TITLE OF INVENTION: SUPERFAMILY (MORPHONS)
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR; TESTA, HURWITZ &
; ADDRESSEE: THIBEAULT, LLP
; STREET: 125 HIGH STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,097A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESQ, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-080
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7000
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..105
; OTHER INFORMATION: /note= "INHIBIN-A SEQUENCE"
US-08-478-097A-22

Query Match 26.9%; Score 46; DB 3; Length 105;
Best Local Similarity 57.1%; Pred. No. 9.2;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 16 gcplhvapyfsypv 29
|||:|:|:|
Db 33 GCGLHIPNLSLPV 46

RESULT 40
US-08-481-377-20
; Sequence 20, Application US/08481377
; Patent No. 5808007
; GENERAL INFORMATION:
; APPLICANT: JOHNS HOPKINS UNIVERSITY
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-3
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: US
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/491,835
; FILING DATE: 23-OCT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00685
; FILING DATE: 12-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: FD3288
; TELECOMMUNICATION INFORMATION:
```

```
COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,377
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00666
; FILING DATE: 12-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WETHERELL, JR. Ph.D., JOHN R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: ED2279 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 121 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: Inhibin alpha
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..121
; OTHER INFORMATION:
US-08-481-377-20

Query Match 26.9%; Score 46; DB 1; Length 121;
Best Local Similarity 57.1%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 16 gcplhvapyfsypv 29
|||:|:|:|
Db 49 GCGLHIPNLSLPV 62

RESULT 41
US-08-491-835-18
; Sequence 18, Application US/08491835
; Patent No. 5821056
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-9
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: US
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/491,835
; FILING DATE: 23-OCT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00685
; FILING DATE: 12-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: FD3288
; TELECOMMUNICATION INFORMATION:
```


TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 121 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
IMMEDIATE SOURCE:
CLONE: Inhibin alpha
FEATURE:
NAME/KEY: Protein
LOCATION: 1..121
US-08-491-835-18

Query Match 25.9%; Score 46; DB 2; Length 121;
Best Local Similarity 57.1%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 16 gcplhvapyfsypv 29
|||:| | | |
Db 49 GCGLHPPNLSLPV 62

RESULT 42
US-09-153-733A-20
Sequence 20, Application US/09153733A
Patent No. 6025475
GENERAL INFORMATION:
APPLICANT: JOHNS HOPKINS UNIVERSITY
TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-3
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
CITY: LOS ANGELES
STATE: CALIFORNIA
COUNTRY: US
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/153,733A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/481,377
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: WETHERELL, JR. Ph.D., JOHN R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: FD2279 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 121 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
IMMEDIATE SOURCE:
CLONE: Inhibin alpha
FEATURE:
NAME/KEY: Protein
LOCATION: 1..121
US-09-153-733A-20

Query Match 26.9%; Score 46; DB 3; Length 121;
Best Local Similarity 57.1%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 16 gcplhvapyfsypv 29
|||:| | | |
Db 49 GCGLHPPNLSLPV 62

RESULT 43
PCT-US94-00666-20
Sequence 20, Application PC/TUS9400666
GENERAL INFORMATION:
APPLICANT: JOHNS HOPKINS UNIVERSITY
TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-3
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
CITY: LOS ANGELES
STATE: CALIFORNIA
COUNTRY: US
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/00666
FILING DATE: 12-JAN-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: WETHERELL, JR. Ph.D., JOHN R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: FD2279 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 121 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
IMMEDIATE SOURCE:
CLONE: Inhibin alpha
FEATURE:
NAME/KEY: Protein
LOCATION: 1..121
PCT-US94-00666-20

Query Match 26.9%; Score 46; DB 4; Length 121;
Best Local Similarity 57.1%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 16 gcplhvapyfsypv 29
|||:| | | |
Db 49 GCGLHPPNLSLPV 62

RESULT 44
PCT-US94-00685-18
Sequence 18, Application PC/TUS9400685
GENERAL INFORMATION:
APPLICANT: THE JOHNS HOPKINS UNIVERSITY
TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-9
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:


```
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: US
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00685
; FILING DATE: 12-JAN-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: FD3288
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 121 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: Inhibin alpha
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..121
; PCT-US94-00685-18

Query Match 26.9%; Score 46; DB 4; Length 121;
Best Local Similarity 57.1%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 16 gcplhvapyfsypv 29
Db 49 GCGLHIPNLSLPV 62

RESULT 45
US-08-581-529B-16
; Sequence 16, Application US/08581529B
; Patent No. 5770444
; GENERAL INFORMATION:
; APPLICANT: Lee, Se-Jin
; APPLICANT: Huynh, Thanh
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-6
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/581,529B
; FILING DATE: 15-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Lisa A. Halle, Ph.D.
```

```
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/082001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: Inhibin-alpha
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..122
; US-08-581-529B-16

Query Match 26.9%; Score 46; DB 1; Length 122;
Best Local Similarity 57.1%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 16 gcplhvapyfsypv 29
Db 50 GCGLHIPNLSLPV 63

RESULT 46
US-08-455-559-22
; Sequence 22, Application US/08455559
; Patent No. 5801014
; GENERAL INFORMATION:
; APPLICANT: Lee, Se-Jin
; APPLICANT: HUYNH, THANH
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-5
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: US
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,559
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/003,144
; FILING DATE: 12-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WETHERELL, JR. PH.D., JOHN R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD2280
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: Inhibin-alpha
```


; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..122
US-08-455-559-22

Query Match 26.9%; Score 46; DB 1; Length 122;
Best Local Similarity 57.1%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 16 gcplhvapyfsypv 29
Db 50 GCGLHIPPNLPLV 63

RESULT 47
US-08-525-596B-26
; Sequence 26, Application US/08525596B
; Patent No. 5827733
; GENERAL INFORMATION:
; APPLICANT: Huynh, Thanh
; APPLICANT: Lee, Se-Jin
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-8
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525.596B
; FILING DATE: 19-SEP-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07762
; FILING DATE: 08-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr., Ph.D, John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07265/075001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5070
; TELEFAX: 619-678-5099
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: Inhibin-alpha
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..122
US-08-525-596B-26

Query Match 26.9%; Score 46; DB 2; Length 122;
Best Local Similarity 57.1%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 16 gcplhvapyfsypv 29
Db 50 GCGLHIPPNLPLV 63

RESULT 48
US-08-581-528A-16
; Sequence 16, Application US/08581528A
; Patent No. 5986058
; GENERAL INFORMATION:
; APPLICANT: Lee, Se-Jin
; APPLICANT: Huynh, Thanh
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-7
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/581,528A
; FILING DATE: 03-Sept-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/089,670
; FILING DATE: 09-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lisa A. Haile, Ph.D.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/081001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: Inhibin-alpha
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..122
US-08-581-528A-16

Query Match 26.9%; Score 46; DB 2; Length 122;
Best Local Similarity 57.1%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 16 gcplhvapyfsypv 29
Db 50 GCGLHIPPNLPLV 63

RESULT 49
PCT-US94-00657-22
; Sequence 22, Application PC/TUS9400657
; GENERAL INFORMATION:
; APPLICANT: SE-JIN LEE
; APPLICANT: HUYNH, THANH
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-5
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA


```
; COUNTRY: US
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00657
; FILING DATE: 1/12/94
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: WETHERELL, JR. PH.D., JOHN R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: FD3256 CIP OF PD2280
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: inhibit-alpha
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..122
; PCT-US94-00657-22

Query Match 26.9%; Score 46; DB 4; Length 122;
Best Local Similarity 57.1%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 16 gcplhvapyfsypv 29
Db 50 GCGLHIPPNLSPV 63

RESULT 50
PCT-US94-07762-16
; Sequence 16, Application PC/TUS9407762
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-6
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07762
; FILING DATE: 08-JUL-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: TUMARKIN, LISA A., PH.D.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: FD2349
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 16:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: Inhibin-alpha
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..122
; PCT-US94-07762-16
```

```
Query Match 26.9%; Score 46; DB 4; Length 122;
Best Local Similarity 57.1%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 16 gcplhvapyfsypv 29
Db 50 GCGLHIPPNLSPV 63
```

```
Search completed: July 14, 2000, 09:31:52
Job time: 2166 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 14, 2000, 08:55:47 ; Search time 53.88 Seconds
(without alignments)
32.645 Million cell updates/sec

Title: BETA-CHAIN
Perfect score: 171
Sequence: 1 vctyrdfiyrtveipgclhvayfsypva 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 168808 seqs, 58629743 residues

Total number of hits satisfying chosen parameters: 168807

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 99%

Listing first 1000 summaries

Database : PIR_63:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
1	159	93.0	138	1	TRTB			thyrotropin beta c
2	156	91.2	138	2	A29479			thyrotropin beta c
3	154	90.1	138	1	TPGB			thyrotropin beta c
4	151	88.3	138	1	TRBOB			thyrotropin beta c
5	110	64.3	147	2	S34148			thyrotropin beta c
6	109	63.7	147	2	A48194			thyrotropin beta c
7	101	59.1	141	1	B60626			gonadotropin beta
8	101	59.1	144	1	UTCAB			gonadotropin beta
9	101	59.1	146	1	S16763			gonadotropin beta
10	99	57.9	142	1	C36179			gonadotropin II be
11	99	57.9	147	2	I50594			gonadotropin II be
12	95	55.6	76	2	IR3048			FSH beta-subunit -
13	95	55.6	141	1	URHUB			lutropin beta chain
14	94	55.0	106	2	S74084			folliotropin beta c
15	93	54.4	130	2	JC4526			folliotropin beta c
16	93	54.4	141	2	JC4527			lutinizing hormon
17	93	54.4	142	2	I51232			gonadotropin II be
18	92	53.8	118	1	FRHOB			folliotropin beta c
19	92	53.8	129	1	FTSHB			folliotropin beta c
20	92	53.8	129	1	FRPGB			folliotropin beta c
21	92	53.8	129	2	A23550			folliotropin beta c
22	92	53.8	130	2	A32893			folliotropin beta c
23	91	53.2	129	1	FRHUB			folliotropin beta c
24	91	53.2	139	2	I52320			testicular luteini
25	91	53.2	141	1	URPFB			lutropin beta chain
26	90	52.6	141	1	URPGB			lutropin beta chain
27	90	52.6	169	1	KTHOB			choriogonadotropin
28	89	52.0	118	2	PN0139			lutropin beta chain
29	89	52.0	119	2	A61465			lutropin beta chain

118	2	PN0141	lutropin beta chain
113	1	S07092	gonadotropin beta
140	2	A48166	gonadotropin II be
142	1	A25800	gonadotropin beta
142	2	I50143	gonadotropin II be
141	1	UTBOB	lutropin beta chain
141	1	UTSHB	lutropin beta chain
141	1	I46949	lutinizing hormon
145	2	I37231	beta-gonadotropin
145	2	I37412	beta-gonadotropin
165	1	KTHOB	choriogonadotropin
112	2	S21196	lutropin beta chain
138	2	S00512	lutropin beta chain
136	2	I50554	gonadotropin II be
165	1	KTBAB	choriogonadotropin
80	2	I65235	testicular luteini
113	2	S07216	gonadotropin I bet
137	2	B36179	gonadotropin I bet
137	2	I51231	gonadotropin I bet
137	2	S34349	gonadotropin I bet
889	2	A35989	modulation protein
262	2	C70372	flagellar hook bas
120	2	I50993	gonadotropin I bet
166	2	I51242	lutinizing hormon
292	2	T00996	hypothetical prote
491	2	T19635	hypothetical prote
128	2	S74085	lutropin beta chain
148	2	E69151	formate hydrogenly
158	2	A61091	lutropin beta chain
159	2	I51373	lutinizing hormon
282	2	T06637	hypothetical prote
711	2	E71673	hypothetical prote
374	2	S69627	hypothetical prote
81	2	A49736	collagen alpha 3(I
220	2	B49736	collagen alpha 3(I
359	2	T35179	oxidoreductase alp
360	2	F70859	hypothetical prote
383	2	T15233	hypothetical prote
448	2	C70375	thiophene and fura
453	2	B70316	DAPA aminotransfer
494	2	A48285	carbamoyl-phosphat
70	2	T31339	collagen alpha 3(I
1645	2	T31339	collagen alpha 3(I
1670	1	GGH03B	insect-selective n
289	2	A45369	hypothetical prote
334	2	T25022	allantoinase homol
446	1	C70016	hypothetical prote
1400	2	T33758	hypothetical prote
341	2	T07145	epoxide hydrolase
341	2	T07152	epoxide hydrolase
366	1	A24248	inhibin alpha chain
382	2	T41256	hypothetical ryano
456	2	JQ1222	hypothetical 50K p
484	1	SYR2ET	glutamate--tRNA li
807	2	S51460	hypothetical prote
1388	2	T34157	hypothetical prote
114	2	I50553	gonadotropin I bet
513	2	T38866	conserved hypoteth
116	2	S37909	hypothetical prote
139	2	JH0615	ribonuclease homol
330	2	T03870	hypothetical prote
348	2	S13323	glucan endo-1,3-be
367	2	F70198	thymidine kinase (
393	2	A49008	paired box transcr
395	1	S70677	probable hydro-ly
398	2	G72333	conserved hypoteth
808	2	T23129	hypothetical prote
3429	2	T13853	hypothetical prote
84	1	JR0152	acrosin/trypsin in
486	1	KRPAAL	ribulose-bisphosph
581	2	G70189	threonine--tRNA li
246	2	I48302	collagen alpha 3(I
308	1	H65050	probable Arf-bindl
324	2	T24242	hypothetical prote

103	44	25.7	327	2	T27352	hypothetical prote	176	41.5	24.3	401	2	D71003	probable aspartate
104	44	25.7	385	2	A55449	aryl-alcohol dehyd	177	41.5	24.3	517	2	A71707	lipopolysaccharide
105	44	25.7	454	2	G65173	thiophene and fura	178	41.5	24.3	633	2	S44795	F0968.3 protein -
106	44	25.7	461	2	G64107	conserved hypotet	179	41.5	24.3	925	2	T29311	hypothetical prote
107	44	25.7	491	1	A69398	M00100 protein hom	180	41.5	24.3	1332	2	T23024	hypothetical prote
108	43.5	25.4	304	2	T75397	metallo-beta-lacta	181	41	24.0	67	2	T17928	hypothetical prote
109	43.5	25.4	388	2	T33908	hypothetical prote	182	41	24.0	116	1	TTUAP	glycoprotein hormo
110	43.5	25.4	398	2	F69863	aspartate aminotra	183	41	24.0	135	2	D64339	hypothetical prote
111	43.5	25.4	576	2	S48354	hypothetical prote	184	41	24.0	152	2	S27791	hypothetical prote
112	43	25.1	134	2	B72851	hypothetical prote	185	41	24.0	166	2	G72383	hypothetical prote
113	43	25.1	140	1	Q0MSIG	hypothetical Ig do	186	41	24.0	176	2	B75015	dtddp-4-keto-l-rham
114	43	25.1	166	2	B72607	hypothetical prote	187	41	24.0	176	2	A71457	hypothetical prote
115	43	25.1	223	2	T09926	cytidine deaminase	188	41	24.0	194	2	T12466	hypothetical prote
116	43	25.1	252	1	J00417	oleoyl-lacvl-carri	189	41	24.0	241	2	E64476	hypothetical prote
117	43	25.1	288	2	F64448	methylviologen-red	190	41	24.0	289	2	S04648	ampR protein - Rho
118	43	25.1	313	2	S08198	cytochrome-c3 hydr	191	41	24.0	306	2	A31800	glucan endo-1,3-be
119	43	25.1	327	2	T23210	hypothetical prote	192	41	24.0	307	2	T09923	cytidine deaminase
120	43	25.1	329	2	A48805	insulin-like growt	193	41	24.0	310	2	JC1434	glucan endo-1,3-be
121	43	25.1	334	2	S20026	beta-glucanase - r	194	41	24.0	312	2	T09443	zinc metalloprotei
122	43	25.1	353	2	B69653	bcsA 5'-region pro	195	41	24.0	315	1	HQDVSB	cytochrome-c3 hydr
123	43	25.1	362	2	T20573	hypothetical prote	196	41	24.0	328	2	T55450	a novel G protein-
124	43	25.1	383	2	S67477	F420-nonreducing h	197	41	24.0	332	2	T02031	1,3-beta-glucanase
125	43	25.1	383	2	S49928	F420-nonreducing h	198	41	24.0	334	2	S05510	glucan endo-1,3-be
126	43	25.1	391	3	T17486	cytochrome P450 hy	199	41	24.0	334	2	D38664	glucan endo-1,3-be
127	43	25.1	441	2	H36789	hypothetical prote	200	41	24.0	335	2	T02088	1,3-beta-glucanase
128	43	25.1	443	1	A46248	dihydroorotate oxi	201	41	24.0	339	2	T24699	hypothetical prote
129	43	25.1	471	2	A39024	collagen alpha 3(I	202	41	24.0	345	1	S56585	L-iditol 2-dehydro
130	43	25.1	492	2	S61442	varlin serine/thre	203	41	24.0	350	2	T25151	hypothetical prote
131	43	25.1	515	2	T25267	hypothetical prote	204	41	24.0	353	2	T33974	hypothetical prote
132	43	25.1	537	2	T25268	embryonic nuclear	205	41	24.0	365	1	S06725	probable hydro-ly
133	43	25.1	539	2	A40581	embryonic nuclear	206	41	24.0	377	2	T40024	probable cytochro
134	43	25.1	602	2	T42401	ubiquitin thiolest	207	41	24.0	380	2	F70399	hydrogenase expres
135	43	25.1	618	2	A71364	probable phosphoen	208	41	24.0	391	2	S23341	paired box transcr
136	43	25.1	651	2	T32875	hypothetical prote	209	41	24.0	403	2	T39846	probable nuclear p
137	43	25.1	653	2	A49722	endoglin precursor	210	41	24.0	411	2	S15296	hypothetical prote
138	43	25.1	719	2	S51739	transcription repr	211	41	24.0	480	1	DCRTA	aromatic-L-amino-a
139	43	25.1	829	1	DJBPT5	DNA-directed DNA p	212	41	24.0	524	1	S40271	virulence factor m
140	43	25.1	926	1	OPPGIT	iodide peroxidase	213	41	24.0	552	1	TVMVT3	transforming prote
141	43	25.1	1299	1	WMBEHS	membrane antigen p	214	41	24.0	656	2	T49299	receptor interacti
142	43	25.1	1435	2	T32930	hypothetical prote	215	41	24.0	660	2	A70362	N-methylhydantoina
143	43	25.1	1616	2	T37183	gene APXI protein	216	41	24.0	704	2	T14614	hypothetical prote
144	43	25.1	1879	2	T19481	hypothetical prote	217	41	24.0	713	1	S46838	hypothetical prote
145	42.5	24.9	114	2	T42672	hypothetical prote	218	41	24.0	782	1	G64157	probable organic s
146	42.5	24.9	389	2	A64346	hypothetical prote	219	41	24.0	811	2	S39901	nwsA protein - Bra
147	42.5	24.9	400	1	I59360	homeotic protein l	220	41	24.0	934	1	G71532	probable threonine
148	42.5	24.9	466	2	T12299	geranylgeranyl hyd	221	41	24.0	996	2	C75498	probable metallopr
149	42.5	24.9	555	2	S69641	hypothetical prote	222	41	24.0	1005	2	T12546	hypothetical prote
150	42.5	24.9	851	1	WMBEU9	gene UL9 protein -	223	41	24.0	1528	2	S13743	DNA strand transfe
151	42.5	24.9	2241	2	S09811	hypothetical prote	224	41	24.0	1964	2	T09059	notch4 - mouse
152	42.5	24.9	3027	2	JQ1917	polyprotein - pars	225	41	24.0	2330	1	RRIWV	genome polyprotein
153	42	24.6	276	2	S26428	hypothetical 31.5K	226	41	24.0	2331	2	S44054	genome polyprotein
154	42	24.6	279	2	T37010	probable DNA-bindi	227	41	24.0	2523	2	F70846	probable PPE prote
155	42	24.6	285	2	D69835	alcohol dehydrogen	228	41	24.0	3014	2	JC5620	polyprotein - hepa
156	42	24.6	289	2	F72637	probable alcohol d	229	40.5	23.7	144	2	S74403	hypothetical prote
157	42	24.6	326	2	T30166	hypothetical prote	230	40.5	23.7	267	2	T33911	hypothetical prote
158	42	24.6	336	2	T25180	hypothetical prote	231	40.5	23.7	327	2	T09027	hypothetical prote
159	42	24.6	357	2	T31829	hypothetical prote	232	40.5	23.7	358	2	A64424	hydrogenase expres
160	42	24.6	415	2	A60086	paired box transcr	233	40.5	23.7	361	2	H75260	conserved hypotet
161	42	24.6	531	2	T23835	hypothetical prote	234	40.5	23.7	365	2	F71437	probable resistanc
162	42	24.6	574	1	S55034	probable 3'-phosph	235	40.5	23.7	367	2	G71076	probable hydrogena
163	42	24.6	650	2	T24739	hypothetical prote	236	40.5	23.7	492	2	G65076	hypothetical prote
164	42	24.6	656	2	B70766	hypothetical prote	237	40.5	23.7	543	1	F64871	oligopeptide-bindi
165	42	24.6	680	2	S26764	major coat protein	238	40.5	23.7	572	1	A53651	probable 3'-phosph
166	42	24.6	722	2	C72040	transcription elon	239	40.5	23.7	816	1	A59003	phosphoinositide 3
167	42	24.6	767	2	F64605	iron(III) dicitrat	240	40.5	23.7	864	2	S71741	DNA ligase (ATP) (
168	42	24.6	1505	2	S26765	genome polyprotein	241	40.5	23.7	925	1	A39216	plasma cell membra
169	42	24.6	2733	2	S15760	genome polyprotein	242	40.5	23.7	1107	1	JQ1658	genome polyprotein
170	41.5	24.3	144	2	G69456	hypothetical prote	243	40.5	23.7	1386	2	T00257	hypothetical prote
171	41.5	24.3	290	2	T07080	shock protein SRC2	244	40.5	23.7	1555	2	S38758	amyl-1,6-glucosid
172	41.5	24.3	325	2	S62592	hypothetical prote	245	40	23.4	103	2	S45783	probable membrane
173	41.5	24.3	325	2	T38139	hypothetical prote	246	40	23.4	143	1	S18241	alpha-amylase inhi
174	41.5	24.3	361	2	E75810	conserved hypotet	247	40	23.4	162	2	C72715	hypothetical prote
175	41.5	24.3	369	2	T27028	hypothetical prote	248	40	23.4	173	2	S24796	hypothetical prote

249	40	23.4	186	2	B75278	conserved hypother	322	39.5	23.1	317	2	T03882	hypothetical prote
250	40	23.4	215	2	A69414	hypothetical prote	323	39.5	23.1	345	2	T16869	hypothetical prote
251	40	23.4	220	2	S76623	O-methyltransferas	324	39.5	23.1	367	2	E75031	hydrogenase expre
252	40	23.4	234	2	G75434	hypothetical prote	325	39.5	23.1	400	2	A69270	molybdenum cofacto
253	40	23.4	240	2	H69535	acetylglutamate ki	326	39.5	23.1	445	2	T32366	hypothetical prote
254	40	23.4	246	2	T05955	1,3-beta-glucanase	327	39.5	23.1	485	2	T27147	hypothetical prote
255	40	23.4	246	2	T05961	1,3-beta-glucanase	328	39.5	23.1	522	1	JT0980	cytochrome P450 52
256	40	23.4	262	2	S76772	hypothetical prote	329	39.5	23.1	579	2	T37248	probable matrix me
257	40	23.4	276	2	H75169	abc transporter PA	330	39.5	23.1	598	2	T32166	hypothetical prote
258	40	23.4	276	2	T10405	probable membrane	331	39.5	23.1	617	2	T16189	hypothetical prote
259	40	23.4	289	2	C69349	conserved hypother	332	39.5	23.1	620	2	JU0273	hemagglutinin - me
260	40	23.4	316	2	T34553	hypothetical prote	333	39.5	23.1	625	2	T30022	hypothetical prote
261	40	23.4	317	2	A36066	trans-activator of	334	39.5	23.1	718	2	T02310	hypothetical prote
262	40	23.4	329	1	E70668	gutg/kpsr family s	335	39.5	23.1	745	2	T23893	hypothetical prote
263	40	23.4	332	2	T09228	malate dehydrogena	336	39.5	23.1	767	2	C71909	iron (III) dicitra
264	40	23.4	332	2	T21399	hypothetical prote	337	39.5	23.1	905	1	A27410	plasma cell membra
265	40	23.4	347	2	S35229	hypD' protein - Br	338	39.5	23.1	915	2	A43802	cellulase (EC 3.2.
266	40	23.4	359	2	G02221	CAGRI protein - hu	339	39.5	23.1	1018	2	A69329	probable isoleucin
267	40	23.4	371	2	S51320	mitogen-activated	340	39.5	23.1	1039	2	S02711	cellulase (EC 3.2.
268	40	23.4	379	2	I39743	hydrogenase homolo	341	39.5	23.1	1241	2	S26373	genome polyprotein
269	40	23.4	379	2	S23443	hypothetical prote	342	39.5	23.1	1242	2	S72350	structural polypro
270	40	23.4	379	2	S29978	hypD protein - Alc	343	39.5	23.1	1242	2	A56605	structural polypro
271	40	23.4	385	2	S32877	hypD protein - Rhi	344	39.5	23.1	1349	2	S51471	killer toxin insen
272	40	23.4	391	2	A40663	paired box transcr	345	39.5	23.1	1530	1	S01393	DNA-directed RNA p
273	40	23.4	401	1	TVFPMH	middle T antigen -	346	39.5	23.1	2294	2	I67630	protein tyrosine p
274	40	23.4	422	2	S71243	3'-phosphoadenosin	347	39.5	23.1	2466	2	I67629	protein tyrosine p
275	40	23.4	437	2	T34211	hypothetical prote	348	39.5	23.1	2485	2	I53483	protein-tyrosine-p
276	40	23.4	447	2	C56146	rfbp protein homol	349	39.5	23.1	2490	1	A54971	protein-tyrosine-p
277	40	23.4	473	2	S18315	ribulose-bisphosph	350	39.5	23.1	2514	1	MNVN2	nonstructural poly
278	40	23.4	475	1	RK2ML	ribulose-bisphosph	351	39.5	23.1	2685	2	T38755	hypothetical prote
279	40	23.4	476	2	S58560	ribulose-bisphosph	352	39	22.8	52	2	S69113	collagen alpha 3(I
280	40	23.4	485	2	T24201	hypothetical prote	353	39	22.8	90	2	D24356	hypothetical prote
281	40	23.4	488	2	S31916	ribulose-bisphosph	354	39	22.8	104	2	D70879	hypothetical prote
282	40	23.4	511	1	B64850	probable virulence	355	39	22.8	108	2	T34677	hypothetical prote
283	40	23.4	550	1	KJHUID	iduronate-2-sulfat	356	39	22.8	133	2	S14728	textilotoxin chain
284	40	23.4	589	2	B70915	probable membrane	357	39	22.8	135	2	T17840	hypothetical prote
285	40	23.4	631	1	JC4602	protein kinase (EC	358	39	22.8	139	2	I32536	T-cell receptor al
286	40	23.4	663	2	T33244	hypothetical prote	359	39	22.8	140	2	C49829	T-cell receptor va
287	40	23.4	668	2	T34317	hypothetical prote	360	39	22.8	143	2	S08466	alpha-amylase inhi
288	40	23.4	668	2	T42211	protein tyrosine p	361	39	22.8	143	2	S13384	alpha-amylase inhi
289	40	23.4	675	2	S16627	penicillin-binding	362	39	22.8	146	2	I48342	phospholipase A2 (
290	40	23.4	682	1	RNRZC1	DNA-directed RNA p	363	39	22.8	149	2	S78524	alpha-amylase inhi
291	40	23.4	683	1	RNRMB1	probable membrane	364	39	22.8	160	1	C70315	2-amino-4-hydroxy-
292	40	23.4	690	2	S35251	hypothetical prote	365	39	22.8	172	2	JE0392	inorganic pyrophos
293	40	23.4	724	2	T19601	hypothetical prote	366	39	22.8	181	2	G65024	hypothetical prote
294	40	23.4	731	2	T09172	probable calcium-a	367	39	22.8	195	2	T15676	hypothetical prote
295	40	23.4	750	2	A56881	prostate-specific	368	39	22.8	198	1	ASLJBT	vif protein - bovi
296	40	23.4	852	2	S41886	DNA repair protein	369	39	22.8	215	2	S57474	GTP-binding protei
297	40	23.4	852	2	T37703	DNA repair protein	370	39	22.8	217	2	T33353	hypothetical prote
298	40	23.4	998	2	I41078	hemolysin - Escher	371	39	22.8	220	2	C75332	hypothetical prote
299	40	23.4	998	2	T42148	hemolysin A toxin	372	39	22.8	230	2	D75518	probable biopolyme
300	40	23.4	998	2	T00227	hemolysin A - Esch	373	39	22.8	238	2	S60081	transcription regu
301	40	23.4	1008	2	S38003	translation elonga	374	39	22.8	246	2	S75275	carboxymethylenebu
302	40	23.4	1055	2	H70951	hypothetical prote	375	39	22.8	277	2	F65014	xanthosine phospho
303	40	23.4	1123	2	D36790	hypothetical prote	376	39	22.8	299	2	H72217	purine nucleoside
304	40	23.4	1168	2	T37496	probable pre-mRNA	377	39	22.8	301	2	T36520	hypothetical prote
305	40	23.4	1220	2	S13057	Ca2+-transporting	378	39	22.8	307	2	T23212	hypothetical prote
306	40	23.4	1266	2	I59314	isoleucine--trNA 1	379	39	22.8	309	2	E36596	nuclear factor I (
307	40	23.4	1305	2	A40879	phospholipase C (E	380	39	22.8	310	2	G02309	UDP-glucuronosyltr
308	40	23.4	1312	1	B40879	phospholipase C (E	381	39	22.8	319	2	B69643	forminoglutamate
309	40	23.4	1355	2	T00075	hypothetical prote	382	39	22.8	323	2	S16318	homeotic protein H
310	40	23.4	1584	2	S57161	hypothetical prote	383	39	22.8	325	2	B72475	probable transcrip
311	40	23.4	1663	1	C3FT	complement C3 prec	384	39	22.8	328	2	S67592	hypothetical prote
312	40	23.4	4574	2	G02520	plectin - human	385	39	22.8	329	2	T24701	hypothetical prote
313	39.5	23.1	106	2	C65031	biphenyl dioxygena	386	39	22.8	330	2	S74255	homeotic protein s
314	39.5	23.1	184	2	T15125	hypothetical prote	387	39	22.8	332	2	T25779	hypothetical prote
315	39.5	23.1	195	2	S44788	D2007.2 protein -	388	39	22.8	339	2	T24059	hypothetical prote
316	39.5	23.1	207	2	A48608	E1 glycoprotein -	389	39	22.8	342	2	T16741	hypothetical prote
317	39.5	23.1	265	2	S38891	connexin 30.3 - ra	390	39	22.8	357	2	PC4293	nuclear factor 1 f
318	39.5	23.1	266	1	A43433	gap junction prote	391	39	22.8	360	2	T21020	hypothetical prote
319	39.5	23.1	273	2	T19359	hypothetical prote	392	39	22.8	364	2	T32589	hypothetical prote
320	39.5	23.1	303	2	JH0401	T-cell-specific tr	393	39	22.8	365	1	Z1BPIK	gene I protein - p
321	39.5	23.1	316	2	T31880	hypothetical prote	394	39	22.8	380	2	S55176	GLG2 protein - yea

395	39	22.8	386	2	S45569	nuclear factor I-A	468	38.5	22.5	146	2	A35493	phospholipase A2 (
396	39	22.8	388	2	S73408	ribose-phosphate p	469	38.5	22.5	253	1	I40444	SpOa activation i
397	39	22.8	400	2	S24739	probable transposa	470	38.5	22.5	259	2	T44433	hypothetical prote
398	39	22.8	400	2	A42727	probable transposa	471	38.5	22.5	261	2	A34476	collagen alpha 2(I
399	39	22.8	444	2	A35071	hydrogen peroxide-	472	38.5	22.5	281	2	T15613	hypothetical prote
400	39	22.8	452	2	A36596	nuclear factor I (473	38.5	22.5	283	2	B38604	poly(3-hydroxyalka
401	39	22.8	455	2	T00856	pectate lyase (EC	474	38.5	22.5	305	2	T32140	hypothetical prote
402	39	22.8	461	2	S45568	nuclear factor I-A	475	38.5	22.5	328	1	B71519	probable gutQ/kpsf
403	39	22.8	464	2	T36256	probable membrane	476	38.5	22.5	347	2	D69373	immunogenic protei
404	39	22.8	467	2	H69109	heat shock protein	477	38.5	22.5	348	2	T08814	1,3-beta-glucanase
405	39	22.8	470	2	H70795	hypothetical prote	478	38.5	22.5	387	2	H69971	spore coat protein
406	39	22.8	472	2	T40755	arginine n-methyl	479	38.5	22.5	398	2	A75128	aspartate aminotra
407	39	22.8	473	1	REBCLT	ribulose-bisphosph	480	38.5	22.5	453	2	A42224	carbamoyl-phosphat
408	39	22.8	483	2	S69894	major DNA-binding	481	38.5	22.5	493	2	E35115	anthranilate synth
409	39	22.8	488	2	A48330	ribulose-bisphosph	482	38.5	22.5	506	1	S54527	aldehyde dehydroge
410	39	22.8	489	2	T24971	hypothetical prote	483	38.5	22.5	506	1	S54615	aldehyde dehydroge
411	39	22.8	490	2	T00404	cytochrome P450 ho	484	38.5	22.5	537	2	I53719	NF-kappa-B transcr
412	39	22.8	493	2	T18789	hypothetical prote	485	38.5	22.5	550	2	A42017	transcription fact
413	39	22.8	498	2	S45567	nuclear factor I-A	486	38.5	22.5	555	2	J01526	interleukin-1 rece
414	39	22.8	505	2	S01300	transcription fact	487	38.5	22.5	562	2	T20188	hypothetical prote
415	39	22.8	506	2	S75370	hypothetical prote	488	38.5	22.5	565	2	T08794	hypothetical prote
416	39	22.8	507	2	T23959	hypothetical prote	489	38.5	22.5	575	2	T22740	hypothetical prote
417	39	22.8	509	2	S45565	nuclear factor I-A	490	38.5	22.5	631	2	JC5803	ring finger protei
418	39	22.8	509	2	JC5428	nuclear factor I f	491	38.5	22.5	638	2	T17554	hypothetical prote
419	39	22.8	511	2	T00864	cytochrome P450 ho	492	38.5	22.5	651	2	A39372	potassium channel
420	39	22.8	514	2	T28855	hypothetical prote	493	38.5	22.5	712	2	T02312	hypothetical prote
421	39	22.8	521	2	S67210	probable membrane	494	38.5	22.5	717	1	VBEB11	glycoprotein H pre
422	39	22.8	522	2	S09996	nuclear factor I-A	495	38.5	22.5	778	2	A60798	platelet glycoprot
423	39	22.8	528	2	T36596	nuclear factor I (496	38.5	22.5	824	2	B53363	desmocollin, type
424	39	22.8	532	2	B36596	nuclear factor I -	497	38.5	22.5	886	2	S53396	telomerase catalyt
425	39	22.8	547	2	E71542	hypothetical prote	498	38.5	22.5	896	2	A53363	desmocollin, type
426	39	22.8	550	1	A34576	crystal protein pr	499	38.5	22.5	968	2	T29532	hypothetical prote
427	39	22.8	574	2	T41207	hypothetical prote	500	38.5	22.5	1110	2	T19673	hypothetical prote
428	39	22.8	601	2	T00119	probable transcrip	501	38.5	22.5	1216	2	A55620	apical endosomal p
429	39	22.8	604	2	T38017	hypothetical prote	502	38.5	22.5	1237	2	B71501	probable DNA pol I
430	39	22.8	614	2	S38802	hypothetical prote	503	38.5	22.5	1758	2	T29350	hypothetical prote
431	39	22.8	637	2	J02039	polyprotein - pars	504	38.5	22.5	1759	2	T29351	collagen alpha 2(I
432	39	22.8	654	2	T41584	hypothetical RNA-b	505	38.5	22.5	2386	2	T39911	rad3 checkpoint pr
433	39	22.8	664	2	S49757	hypothetical prote	506	38.5	22.5	2491	1	A28372	insulin-like growt
434	39	22.8	716	1	PC4225	replication licens	507	38.5	22.5	4273	2	C69679	polyketide synthas
435	39	22.8	752	2	T09082	telomere-associate	508	38.5	22.5	4544	1	S02392	alpha-2-macroglobu
436	39	22.8	755	2	T00066	hypothetical prote	509	38.5	22.5	4545	1	S25111	alpha-2-macroglobu
437	39	22.8	757	2	T09081	telomere-associate	510	38.5	22.5	80	2	PC4291	LiMk2a protein - h
438	39	22.8	823	2	S44873	ZC21.2 protein - C	511	38	22.2	84	2	T17637	hypothetical prote
439	39	22.8	845	2	JC5256	adipocyte transcr	512	38	22.2	86	2	S44154	hypothetical prote
440	39	22.8	933	1	OPHUIT	iodide peroxidase	513	38	22.2	112	1	WMBEH5	UL35 protein - hum
441	39	22.8	955	2	T33040	hypothetical prote	514	38	22.2	129	2	T06188	probable DNA-direc
442	39	22.8	971	2	T24866	hypothetical prote	515	38	22.2	132	2	B27261	hypothetical prote
443	39	22.8	1135	1	GNVUHV	M polyprotein - Ha	516	38	22.2	154	2	T37489	40S ribosomal prot
444	39	22.8	1135	1	GNVUHV	M polyprotein - Ha	517	38	22.2	155	1	T1ZM1	trypsin/factor XII
445	39	22.8	1135	2	JQ50605	M polyprotein - Ha	518	38	22.2	175	1	WCHCS	coagulogen - horse
446	39	22.8	1224	2	T14007	microtubule-associ	519	38	22.2	181	2	A05080	kappa-casein precu
447	39	22.8	1239	2	G02750	DNA-directed DNA p	520	38	22.2	186	2	A64707	peptidyl-tRNA hydr
448	39	22.8	1333	1	XOHUHD	xanthine dehydroge	521	38	22.2	194	2	T33719	hypothetical prote
449	39	22.8	1339	2	JC5080	DNA-directed DNA p	522	38	22.2	203	2	S59279	hypothetical prote
450	39	22.8	1357	2	T16860	hypothetical prote	523	38	22.2	204	2	T41513	probable proteasom
451	39	22.8	1371	2	A33837	insulin-like growt	524	38	22.2	207	2	S45226	yads protein - Esc
452	39	22.8	1390	2	T18883	hypothetical prote	525	38	22.2	213	2	E70591	hypothetical prote
453	39	22.8	1775	2	A31893	collagen alpha 1(I	526	38	22.2	219	2	S27610	hypothetical prote
454	39	22.8	1786	2	A57282	ankyrin-related pr	527	38	22.2	220	1	E85000	NADH dehydrogenase
455	39	22.8	1809	2	T15345	ankyrin-related un	528	38	22.2	223	2	S77586	hypothetical prote
456	39	22.8	1815	2	T15346	elegans ankyrin-re	529	38	22.2	225	2	JX0259	C-reactive protein
457	39	22.8	1867	2	T15344	ankyrin-related un	530	38	22.2	228	2	T39386	probable GI-S-spec
458	39	22.8	2039	2	T15347	ankyrin-related un	531	38	22.2	244	2	T24631	hypothetical prote
459	39	22.8	2529	2	A56923	transcription fact	532	38	22.2	253	2	JN0771	hypothetical 28K p
460	39	22.8	2559	2	C69681	peptide synthetase	533	38	22.2	261	2	A40579	trans-activating t
461	39	22.8	2578	2	A56922	transcription fact	534	38	22.2	265	2	B48151	sperm tail protein
462	39	22.8	2647	2	A37098	gelation factor AB	535	38	22.2	272	2	B75414	molymdenum cofacto
463	39	22.8	3712	2	S18253	laminin alpha-1 ch	536	38	22.2	273	2	B69883	conserved hypothet
464	39	22.8	7576	2	T17428	FK506 polyketide s	537	38	22.2	273	2	T28973	hypothetical prote
465	38.5	22.5	128	2	S69856	hypothetical prote	538	38	22.2	285	2	G02240	UDP-glucuronosyltr
466	38.5	22.5	146	2	A33394	phospholipase A2 (539	38	22.2	291	2	S33209	extensin-like prot
467	38.5	22.5					540	38	22.2	293	1	S00112	glycine N-methyltr

541	38	22.2	293	2	F09920	cytidine deaminase	614	38	22.2	786	2	T08664	Toll protein-like
542	38	22.2	307	2	D64206	ribose-phosphate p	615	38	22.2	819	1	TVCHFG	fibroblast growth
543	38	22.2	320	2	T40240	dimethylase - f1ss	616	38	22.2	821	1	CIHUB3	calpain [EC 3.4.22
544	38	22.2	308	2	T24453	hypothetical prote	617	38	22.2	823	2	T16758	hypothetical prote
545	38	22.2	309	2	C72589	hypothetical prote	618	38	22.2	892	2	A41697	nitrate assimilati
546	38	22.2	312	2	JC1438	glucan endo-1,3-be	619	38	22.2	944	2	T38130	probable helicase
547	38	22.2	312	2	T32379	hypothetical prote	620	38	22.2	956	2	H75536	2-oxoglutarate deh
548	38	22.2	316	2	S46237	glucan endo-1,3-be	621	38	22.2	958	2	S41013	hypothetical prote
549	38	22.2	316	2	T23368	hypothetical prote	622	38	22.2	1032	2	T23164	hypothetical prote
550	38	22.2	320	2	T21401	hypothetical prote	623	38	22.2	1035	2	T23165	hypothetical prote
551	38	22.2	321	2	T25004	hypothetical prote	624	38	22.2	1042	2	S76045	hypothetical prote
552	38	22.2	325	2	S75095	hypothetical prote	625	38	22.2	1051	1	JW0051	serine/threonine-s
553	38	22.2	328	2	T20582	hypothetical prote	626	38	22.2	1069	2	S27922	nuclear antigen EB
554	38	22.2	333	2	S15309	hypothetical prote	627	38	22.2	1071	2	S44798	F09G8.5 protein -
555	38	22.2	334	2	T02210	1,3-beta-glucanase	628	38	22.2	1080	2	S30876	hypothetical prote
556	38	22.2	334	2	T19955	hypothetical prote	629	38	22.2	1200	2	T17404	hyalin - sea urchi
557	38	22.2	334	2	A48151	sperm tail protein	630	38	22.2	1221	2	T23472	hypothetical prote
558	38	22.2	336	2	T19757	hypothetical prote	631	38	22.2	1233	1	G71612	novel serine/threo
559	38	22.2	337	1	S56389	hypothetical 37.7k	632	38	22.2	1254	1	A32686	DNA-directed DNA p
560	38	22.2	337	2	T09925	cytidine deaminase	633	38	22.2	1646	2	T20740	hypothetical prote
561	38	22.2	342	1	G70460	conserved hypothet	634	38	22.2	1763	2	S16366	collagen alpha 2(I
562	38	22.2	345	2	G69492	hypothetical prote	635	38	22.2	1827	1	UUHU	sucrose alpha-gluc
563	38	22.2	347	2	A70173	femaA protein (fema	636	38	22.2	1973	2	T18686	hypothetical prote
564	38	22.2	349	2	T25236	hypothetical prote	637	38	22.2	2164	1	GNNY89	genome polyprotein
565	38	22.2	362	2	D72424	oligopeptide ABC t	638	38	22.2	2305	2	T15571	hypothetical prote
566	38	22.2	362	2	A71369	conserved hypothet	639	38	22.2	2470	2	S50726	cation-independent
567	38	22.2	364	2	S22394	retuin precursor -	640	38	22.2	2970	2	T08839	polyprotein - marm
568	38	22.2	366	2	T05807	hypothetical prote	641	38	22.2	3924	2	S37431	ankyrin 2, neuroma
569	38	22.2	366	2	G71067	hypothetical prote	642	38	22.2	9376	2	T14593	streptomycin synth
570	38	22.2	367	2	D72300	oligopeptide ABC t	643	37.5	21.9	150	2	S22328	gene D protein - p
571	38	22.2	374	2	S74470	hydrogenase isoenz	644	37.5	21.9	152	1	ZDBP74	gene D protein - p
572	38	22.2	380	2	JL0053	upDglucose--hexose	645	37.5	21.9	152	2	ZDBP64	gene D protein - p
573	38	22.2	383	1	DHHU	Ig delta chain c r	646	37.5	21.9	152	2	JS0454	gene D protein - p
574	38	22.2	394	1	S43915	maltose permease -	647	37.5	21.9	168	2	S10029	alpha-amylase inhi
575	38	22.2	397	2	B69760	conserved hypothet	648	37.5	21.9	168	2	S16031	hypothetical prote
576	38	22.2	437	2	C70960	hypothetical prote	649	37.5	21.9	176	2	D71146	pfis protein - Esch
577	38	22.2	451	2	E72398	hypothetical prote	650	37.5	21.9	232	2	S45227	gene 2 protein - h
578	38	22.2	454	2	T16429	hypothetical prote	651	37.5	21.9	238	1	WZBE2	hypothetical prote
579	38	22.2	458	2	S36252	constitutive potas	652	37.5	21.9	247	2	T22535	pseudouridyate sy
580	38	22.2	458	2	S48001	sapG protein - Sal	653	37.5	21.9	280	2	D75291	dihydrodipicolinat
581	38	22.2	460	2	JC4214	translation elonga	654	37.5	21.9	298	2	A64058	cytidine deaminase
582	38	22.2	468	2	S52495	acid phosphatase h	655	37.5	21.9	298	2	T09921	purine nucleoside
583	38	22.2	471	1	RKKRL2	ribulose-bisphosph	656	37.5	21.9	305	2	D75306	hypothetical prote
584	38	22.2	476	2	A70477	glutamate synthase	657	37.5	21.9	308	2	T31890	proteoglycan-Ib -
585	38	22.2	480	1	DEGPA	Ig delta chain (WI	658	37.5	21.9	316	2	A41781	iron (iii) abc tra
586	38	22.2	509	2	S17597	phosphoprotein pho	659	37.5	21.9	331	2	G75035	mannose-1-phosphat
587	38	22.2	530	2	S46322	UDP glucuronosyltr	660	37.5	21.9	336	2	H72303	iron stress-induce
588	38	22.2	530	2	JC5856	glucuronosyltransf	661	37.5	21.9	342	2	A30189	photosystem II chl
589	38	22.2	530	2	S17512	cytochrome-c oxida	662	37.5	21.9	344	2	S42648	hypothetical prote
590	38	22.2	538	2	S36424	protein kinase (EC	663	37.5	21.9	344	2	T05437	N-acetyl-gamma-glu
591	38	22.2	543	1	JC4070	hypothetical prote	664	37.5	21.9	351	1	G76613	probable hemin ABC
592	38	22.2	544	2	S46099	pollen-specific pr	665	37.5	21.9	354	2	G75619	vbG0 protein - Esc
593	38	22.2	547	2	T02910	hypothetical prote	666	37.5	21.9	362	2	C64807	hydrogenase expres
594	38	22.2	555	2	E71420	hypothetical prote	667	37.5	21.9	370	2	F71883	hypothetical prote
595	38	22.2	562	2	G70002	hypothetical prote	668	37.5	21.9	372	2	T24738	transforming prote
596	38	22.2	565	1	VGNZSV	cell fusion glycop	669	37.5	21.9	380	1	TVMSF	omega-6 desaturase
597	38	22.2	565	1	VGNZSH	cell fusion glycop	670	37.5	21.9	387	2	T07687	hypothetical prote
598	38	22.2	565	1	VGNZFS	cell fusion glycop	671	37.5	21.9	411	2	T19728	NADH dehydrogenase
599	38	22.2	565	2	T22649	hypothetical prote	672	37.5	21.9	432	2	G71683	complement factor
600	38	22.2	576	2	T36902	probable dihydroxy	673	37.5	21.9	449	1	NBHHS	probable seryl-TNN
601	38	22.2	584	2	S55106	probable membrane	674	37.5	21.9	463	2	B72500	interstitial colla
602	38	22.2	585	2	C70330	conserved hypothet	675	37.5	21.9	468	1	KCRBI	alpha-amylase (EC
603	38	22.2	587	2	T32546	hypothetical prote	676	37.5	21.9	495	1	A45738	6-aminohexanoate-c
604	38	22.2	598	2	E35255	2,4-dichlorophenol	677	37.5	21.9	506	2	S75789	DNA-directed RNA p
605	38	22.2	598	2	T20769	hypothetical prote	678	37.5	21.9	511	2	S73104	conserved hypothet
606	38	22.2	612	2	T34290	hypothetical prote	679	37.5	21.9	529	2	C69280	protein kinase cdr
607	38	22.2	632	2	C71327	probable phosphogl	680	37.5	21.9	593	1	KIZPMN	hypothetical prote
608	38	22.2	656	1	QOBT2	UL25 protein - hum	681	37.5	21.9	682	2	T21420	transferrin precur
609	38	22.2	668	1	S74619	hypothetical prote	682	37.5	21.9	694	1	TFRBP	nuclear autoantige
610	38	22.2	688	2	T31633	hypothetical prote	683	37.5	21.9	713	2	JC2522	norbin - rat
611	38	22.2	706	2	T25897	hypothetical prote	684	37.5	21.9	729	2	JC5812	probable protein-t
612	38	22.2	710	2	B71374	probable soluble 1	685	37.5	21.9	807	1	S44538	hypothetical prote
613	38	22.2	769	1	IJHULM	leukocyte adhesion	686	37.5	21.9	835	2	E72305	

687	37.5	21.9	970	2	A41944	mitotic control pr	760	37	21.6	366	2	H70417	cytochrome oxidase
688	37.5	21.9	1032	1	G7BP74	baseplate protein	761	37	21.6	366	2	H69303	iron (III) ABC tra
689	37.5	21.9	1081	2	S15040	pleiotropic drug r	762	37	21.6	366	2	A50777	iron (III) abc tra
690	37.5	21.9	1091	2	S57112	JSNI protein - yea	763	37	21.6	368	2	S39478	glutamate--ammonia
691	37.5	21.9	1224	1	ERHDAH	coatomer complex a	764	37	21.6	370	2	B64632	hydrogenase expres
692	37.5	21.9	1231	1	NBHUH	complement factor	765	37	21.6	371	2	A69421	hydrogenase expres
693	37.5	21.9	1335	2	T30211	autolysin E - Stap	766	37	21.6	372	2	C65086	hypothetical prote
694	37.5	21.9	1487	2	T02850	probable membrane	767	37	21.6	372	2	T03385	naringenin 3-dioxy
695	37.5	21.9	1515	2	A40203	4-alpha-glucanotra	768	37	21.6	373	2	S15200	hydrogenase expres
696	37.5	21.9	1626	2	A75613	hypothetical prote	769	37	21.6	377	2	S14647	naringenin 3-dioxy
697	37.5	21.9	1707	2	S77908	hypothetical prote	770	37	21.6	378	2	JG5689	CIRCE binding repr
698	37.5	21.9	1978	2	T07081	acetyl-CoA carboxy	771	37	21.6	390	2	F69447	conserved hypothet
699	37.5	21.9	2482	2	I48922	cation-independent	772	37	21.6	401	2	T05025	heat shock transcr
700	37.5	21.9	2483	1	A49617	insulin-like growt	773	37	21.6	402	2	S75082	pancreatic metab
701	37	21.6	118	2	A30505	ig lambda chain C	774	37	21.6	403	2	B35401	cytochrome P450 10
702	37	21.6	120	2	C69947	phage-related prot	775	37	21.6	416	2	S76329	probable phosphoes
703	37	21.6	122	1	WILAI	alpha-amylase/tyrp	776	37	21.6	426	2	A35641	5-aminoimidazole r
704	37	21.6	127	2	A72655	hypothetical prote	777	37	21.6	428	2	A75423	multidrug resistan
705	37	21.6	132	2	S03224	hypothetical prote	778	37	21.6	445	2	A70732	hypothetical prote
706	37	21.6	137	2	S33547	hypothetical prote	779	37	21.6	448	2	A51175	hypothetical prote
707	37	21.6	145	2	G72451	hypothetical prote	780	37	21.6	450	2	T40631	probable membrane
708	37	21.6	151	2	T21828	hypothetical prote	781	37	21.6	450	2	T40337	probable saccharop
709	37	21.6	153	2	D64364	formate hydrogenly	782	37	21.6	453	2	H70333	conserved hypothet
710	37	21.6	160	2	B53814	p20 protein - huma	783	37	21.6	459	2	S17742	prochlorophyllid
711	37	21.6	162	2	A53814	p20 protein - rat	784	37	21.6	461	2	S13806	translation elonga
712	37	21.6	171	2	T15894	hypothetical prote	785	37	21.6	461	2	JH0530	translation elonga
713	37	21.6	180	2	S39775	alpha-s2-casein fo	786	37	21.6	462	2	T01732	UTP-glucose glucos
714	37	21.6	181	2	T20323	hypothetical prote	787	37	21.6	462	2	A60491	translation elonga
715	37	21.6	182	2	A65073	hypothetical prote	788	37	21.6	462	2	S50143	translation elonga
716	37	21.6	182	2	T15754	hypothetical prote	789	37	21.6	462	2	I50226	translation elonga
717	37	21.6	182	2	T15753	hypothetical prote	790	37	21.6	465	1	W2EBE4	51.3K capsid prote
718	37	21.6	186	1	C71206	hypothetical prote	791	37	21.6	467	2	T07565	probable protochlo
719	37	21.6	188	2	I59421	mast cell function	792	37	21.6	472	2	D71076	probable glutamate
720	37	21.6	194	2	F64860	hypothetical prote	793	37	21.6	473	1	A38874	protein-tyrosine-p
721	37	21.6	208	2	T17329	hypothetical prote	794	37	21.6	473	2	S59984	hypb protein - Rho
722	37	21.6	210	2	A49180	glutathione transf	795	37	21.6	473	2	J01936	xylan 1,4-beta-xy
723	37	21.6	213	2	T36895	probable 3-methyla	796	37	21.6	473	2	J50770	xylosidase (EC 3.2
724	37	21.6	222	2	A36730	nutg protein - kie	797	37	21.6	487	2	T21384	hypothetical prote
725	37	21.6	224	2	S55081	hypothetical prote	798	37	21.6	492	2	T23502	hypothetical prote
726	37	21.6	238	2	T05959	1,3-beta-glucanase	799	37	21.6	500	1	S22530	carboxypeptidase C
727	37	21.6	238	2	T05962	1,3-beta-glucanase	800	37	21.6	502	2	S43123	protein-tyrosine-p
728	37	21.6	239	2	C72771	hypothetical prote	801	37	21.6	504	2	T10698	legumin-like prote
729	37	21.6	247	2	S05494	trypsin (EC 3.4.21	802	37	21.6	509	2	C89491	probable acid-CoA
730	37	21.6	248	2	G69113	hypothetical prote	803	37	21.6	510	2	A52507	glycerol-3-phospha
731	37	21.6	250	2	G72495	probable polysulfi	804	37	21.6	518	2	S19514	probable membrane
732	37	21.6	255	2	T06359	1,3-beta-glucanase	805	37	21.6	534	2	T32020	hypothetical prote
733	37	21.6	255	2	E72522	hypothetical prote	806	37	21.6	548	2	T05670	pollen-specific pr
734	37	21.6	260	1	XMEBFC	flagellar basal bo	807	37	21.6	555	2	I51671	Wee1A kinase - Aif
735	37	21.6	260	2	C64851	flagellar basal bo	808	37	21.6	556	2	A44441	B-cell antigen CD1
736	37	21.6	261	2	B55558	thyrotroph embryon	809	37	21.6	562	2	G02426	interleukin-1 rece
737	37	21.6	264	1	B69413	conserved hypothet	810	37	21.6	564	2	S16304	benzoyl/methotrex
738	37	21.6	264	2	C71259	probable flagellar	811	37	21.6	565	2	S20534	cytochrome-c oxida
739	37	21.6	286	2	A69168	hypothetical prote	812	37	21.6	565	2	T10695	legumin-like prote
740	37	21.6	292	2	T32805	hypothetical prote	813	37	21.6	574	2	T23668	hypothetical prote
741	37	21.6	293	2	J01701	hypothetical 3.2K	814	37	21.6	577	2	F72122	hypothetical prote
742	37	21.6	294	2	T36949	conserved hypothet	815	37	21.6	583	2	A45381	phytoene dehydroge
743	37	21.6	297	2	T18637	hypothetical prote	816	37	21.6	608	2	C72405	hypothetical prote
744	37	21.6	300	2	T37759	c-5 sterol desatur	817	37	21.6	622	2	T02244	probable DNA repai
745	37	21.6	303	2	B43451	Na+/K+-exchangin	818	37	21.6	629	2	T19563	hypothetical prote
746	37	21.6	303	2	G02360	thyrotroph embryon	819	37	21.6	640	1	Q1CHGM	phosphoenolpyruvat
747	37	21.6	306	2	S77133	homoserine kinase	820	37	21.6	650	2	A54065	sodium transport p
748	37	21.6	308	2	C30315	coenzyme-M-7-merca	821	37	21.6	654	2	T10521	beta-glucosidase (
749	37	21.6	308	2	D30315	methyl viologen-re	822	37	21.6	690	2	H75469	conserved hypothet
750	37	21.6	311	1	DCLBHP	histidine decarbox	823	37	21.6	695	2	S76771	arginine decarboxy
751	37	21.6	312	2	T32211	hypothetical prote	824	37	21.6	703	2	H72034	thio:disulfide int
752	37	21.6	320	2	F75082	hypothetical prote	825	37	21.6	726	2	S74514	hypothetical prote
753	37	21.6	335	2	S64352	hypothetical prote	826	37	21.6	728	1	TVHUSK	transforming prote
754	37	21.6	354	2	T20169	hypothetical prote	827	37	21.6	733	1	S56767	replication licens
755	37	21.6	357	2	S38338	naringenin 3-dioxy	828	37	21.6	734	1	I38080	probable serine/th
756	37	21.6	358	1	J00805	hydrogenase (EC 1.	829	37	21.6	737	1	S64767	hypothetical prote
757	37	21.6	359	2	A42462	acetoin catabolism	830	37	21.6	738	2	D70680	phenylalanine--trn
758	37	21.6	359	2	T21840	hypothetical prote	831	37	21.6	788	2	A72330	probable ponA' pro
759	37	21.6	365	2	S08092	gene I protein - p	832	37	21.6	810	2	C70791	

833	37	21.6	816	2	S46268	atatin-1 - human	906	36.5	21.3	379	2	T32695	hypothetical prote
834	37	21.6	825	1	A60386	interleukin-4 rece	907	36.5	21.3	381	2	C70036	capsular polysacch
835	37	21.6	840	2	C70131	leucine--tRNA liga	908	36.5	21.3	381	2	A54415	transcription fact
836	37	21.6	861	2	S73648	isoleucine--tRNA l	909	36.5	21.3	382	2	I39695	beta-lactamase (EC
837	37	21.6	865	2	F64734	aconitate hydrat	910	36.5	21.3	386	2	C70505	probable acyl-coa
838	37	21.6	873	1	S53828	cytochrome-c oxida	911	36.5	21.3	397	2	I58099	gene P2X3 protein
839	37	21.6	887	2	T25358	hypothetical prote	912	36.5	21.3	397	2	S60334	purinoceptor P2X -
840	37	21.6	896	2	B43817	transforming prote	913	36.5	21.3	407	2	S69286	geranylgeranyl hyd
841	37	21.6	946	2	T08913	cell division prot	914	36.5	21.3	411	2	C30091	capA protein - Bac
842	37	21.6	947	2	T23107	hypothetical prote	915	36.5	21.3	415	2	T40535	probable arginine-
843	37	21.6	953	2	D71645	exonuclease ABC c	916	36.5	21.3	437	2	S60957	transcription modu
844	37	21.6	990	1	SNEFIN	insulinase (EC 3.4	917	36.5	21.3	446	2	C70783	hypothetical prote
845	37	21.6	1020	2	A29355	fibronectin - chic	918	36.5	21.3	450	2	B40392	alpha-2-adrenergic
846	37	21.6	1034	2	S63536	aminomethyltransfe	919	36.5	21.3	450	2	JH0190	alpha-2-adrenergic
847	37	21.6	1034	2	T22166	hypothetical prote	920	36.5	21.3	460	2	B48057	farnesyl-diphospha
848	37	21.6	1035	2	T07826	aminomethyltransfe	921	36.5	21.3	460	2	T40581	hypothetical prote
849	37	21.6	1037	2	S63535	aminomethyltransfe	922	36.5	21.3	463	2	T16503	hypothetical prote
850	37	21.6	1037	2	S40216	p protein - Flaver	923	36.5	21.3	467	2	T32292	hypothetical prote
851	37	21.6	1073	2	B36429	integrin alpha-6 c	924	36.5	21.3	481	2	A27626	sucrose phosphoryl
852	37	21.6	1088	2	T13950	1-phosphatidylinos	925	36.5	21.3	483	2	G75077	hypothetical prote
853	37	21.6	1091	2	A41543	integrin alpha-6 c	926	36.5	21.3	487	1	BWSOGM	gtfA protein - Str
854	37	21.6	1093	2	A41251	valine--tRNA ligas	927	36.5	21.3	488	2	S78236	ribulose-bisphosph
855	37	21.6	1143	2	S46122	SNF2 protein homol	928	36.5	21.3	506	2	JC5678	minor core protein
856	37	21.6	1143	2	T22952	hypothetical prote	929	36.5	21.3	507	1	Q0BE41	BGLF1 protein - hu
857	37	21.6	1155	2	G64332	delta-endotoxin -	930	36.5	21.3	508	2	A33378	fascilin III prec
858	37	21.6	1434	2	T12322	hypothetical prote	931	36.5	21.3	520	1	S57337	trichodiene oxgen
859	37	21.6	1459	2	G74088	hypothetical prote	932	36.5	21.3	523	2	A31948	carnitine octanoyl
860	37	21.6	1276	2	T27859	ankyrin - fruit fl	933	36.5	21.3	532	2	H64110	L-lactate permease
861	37	21.6	1319	2	S49951	SSM4 protein - yea	934	36.5	21.3	532	2	A37932	nuclear factor kap
862	37	21.6	1357	2	T29265	hypothetical prote	935	36.5	21.3	549	2	S48493	malate synthase (E
863	37	21.6	1385	2	T18213	delta-endotoxin -	936	36.5	21.3	554	2	JE0338	Frizzled-2 protein
864	37	21.6	1434	2	G71232	hypothetical prote	937	36.5	21.3	565	2	S60426	probable membrane
865	37	21.6	1459	2	G74088	hypothetical prote	938	36.5	21.3	570	2	S61650	hypothetical prote
866	37	21.6	1549	2	T13940	ankyrin - fruit fl	939	36.5	21.3	572	2	JG0181	XiLL2 protein - hu
867	37	21.6	1570	2	T38792	probable protein t	940	36.5	21.3	575	2	T40676	hypothetical prote
868	37	21.6	1608	2	T17201	adenylate cyclase	941	36.5	21.3	598	2	S76959	GTP-binding membra
869	37	21.6	1635	2	T14075	chitinase (EC 3.2.	942	36.5	21.3	603	2	S60025	carnitine octanoyl
870	37	21.6	1811	2	T39252	probable protein t	943	36.5	21.3	612	2	G69114	indolepyruvate oxi
871	36.5	21.3	102	2	E72501	hypothetical prote	944	36.5	21.3	618	2	I39196	amiloride sensitiv
872	36.5	21.3	109	2	E75294	hypothetical prote	945	36.5	21.3	634	2	S41788	endo-1,4-beta-xyla
873	36.5	21.3	110	1	IPDG	insulin precursor	946	36.5	21.3	684	2	B71376	hypothetical prote
874	36.5	21.3	111	2	T07656	probable resistanc	947	36.5	21.3	722	2	A54411	amine oxidase (cop
875	36.5	21.3	125	2	A05160	hypothetical prote	948	36.5	21.3	762	2	C64461	hypothetical prote
876	36.5	21.3	135	2	H75451	gamma-carboxymucon	949	36.5	21.3	821	2	A34347	translation elonga
877	36.5	21.3	144	2	S39560	alpha-amylose inhi	950	36.5	21.3	830	2	T19283	hypothetical prote
878	36.5	21.3	145	2	S10027	alpha-amylose inhi	951	36.5	21.3	835	1	QJBEHA	BSLF1 protein - sa
879	36.5	21.3	167	2	T36290	probable integral	952	36.5	21.3	836	1	JQVLUD	DNA-directed DNA p
880	36.5	21.3	176	1	PWEC	inorganic pyrophos	953	36.5	21.3	848	2	T00788	protein kinase hom
881	36.5	21.3	181	2	S38460	plastoquinol--plas	954	36.5	21.3	880	2	T18771	probable chloride
882	36.5	21.3	188	2	D53304	transfer protein D	955	36.5	21.3	881	2	T31818	hypothetical prote
883	36.5	21.3	199	2	T27239	hypothetical prote	956	36.5	21.3	887	1	RDHYE	hydroxymethylgluta
884	36.5	21.3	210	2	T22344	hypothetical prote	957	36.5	21.3	887	1	A23586	hydroxymethylgluta
885	36.5	21.3	213	2	S55129	transcription fact	958	36.5	21.3	888	1	RDHUE	hydroxymethylgluta
886	36.5	21.3	217	2	A70672	hypothetical prote	959	36.5	21.3	914	2	T00757	hypothetical prote
887	36.5	21.3	257	2	T18488	hypothetical prote	960	36.5	21.3	926	2	T39664	chitin synthase 2
888	36.5	21.3	267	2	T20096	hypothetical prote	961	36.5	21.3	947	2	T20512	hypothetical prote
889	36.5	21.3	269	2	T05895	xyloglucan endo-1,	962	36.5	21.3	963	2	B70524	probable PPE prote
890	36.5	21.3	279	2	T09870	probable endo-xyla	963	36.5	21.3	997	2	S63064	probable membrane
891	36.5	21.3	293	2	T06027	xyloglucan endo-1,	964	36.5	21.3	1145	2	SL3643	PRP22 protein - ye
892	36.5	21.3	302	2	F64883	probable transcrip	965	36.5	21.3	1331	1	XORDDH	xanthine dehydroge
893	36.5	21.3	304	2	H75081	phosphate abc tran	966	36.5	21.3	1335	1	XOMSDH	xanthine dehydroge
894	36.5	21.3	317	2	T20018	hypothetical prote	967	36.5	21.3	1358	1	XOCHDH	xanthine dehydroge
895	36.5	21.3	321	2	S55640	hypothetical prote	968	36.5	21.3	1696	2	T27447	hypothetical prote
896	36.5	21.3	322	1	B70434	polysialic acid ca	969	36.5	21.3	1752	2	A45407	collagen alpha 3(I
897	36.5	21.3	331	2	T26437	hypothetical prote	970	36.5	21.3	1834	1	JDMU1	DNA-directed RNA p
898	36.5	21.3	348	1	BVRCMB	mrc protein - Esc	971	36.5	21.3	1834	1	A30788	mannose 6-phosphat
899	36.5	21.3	348	2	D69009	hydrogenase expres	972	36.5	21.3	2499	1	T00252	MEGFI protein - ra
900	36.5	21.3	350	2	T29409	hypothetical prote	973	36	21.1	4351	2	H72124	hypothetical prote
901	36.5	21.3	351	2	J02166	spindle body prote	974	36	21.1	52	2	B61384	Cowper's gland muc
902	36.5	21.3	352	2	JC5185	enhancing factor p	975	36	21.1	61	2	G42529	B-ORF-F protein -
903	36.5	21.3	356	2	T13430	hypothetical prote	976	36	21.1	75	2	J01831	hypothetical prote
904	36.5	21.3	362	2	D70603	probable lipoprote	977	36	21.1	81	1	S29820	acrosin/trypsin in
905	36.5	21.3	377	2	I50147	3beta-hydroxy-Delt	978	36	21.1	102	2	B49056	T-cell receptor al

979 36 21.1 103 2 S40134
980 36 21.1 109 1 L7RB
981 36 21.1 109 2 JW0100
982 36 21.1 111 2 S51032
983 36 21.1 120 2 S53062
984 36 21.1 120 2 S43030
985 36 21.1 120 2 S46185
986 36 21.1 125 2 E70332
987 36 21.1 128 2 C48552
988 36 21.1 138 1 F64371
989 36 21.1 138 2 PC2226
990 36 21.1 139 2 I50677
991 36 21.1 140 2 S67666
992 36 21.1 141 1 HANER
993 36 21.1 145 2 S13376
994 36 21.1 150 2 S72361
995 36 21.1 151 2 S30448
996 36 21.1 151 2 S30453
997 36 21.1 151 2 S30452
998 36 21.1 151 2 S30450
999 36 21.1 151 2 S30451

T-cell receptor V-
Ig lambda chain C
azaarene carbazole
subtilisin-chymotr
glycoprotein hormo
hypothetical prote
hypothetical prote
conserved hypotet
hypothetical prote
hypothetical prote
peptide LAIE 4 - p
ribonuclease A/ang
probable membrane
hemoglobin alpha c
CM2 protein - duru
pancreatic ribonuc
env protein - huma
env protein - huma
env protein - huma
env protein - huma
env protein - huma

ALIGNMENTS

RESULT 1
TTRFB
Thyrotropin beta chain precursor - rat
N:Alternate names: TSH
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 18-Jun-1999
C:Accession: A90954; A90946; A26706; I52206; A01492
R:Croyle, M.L.; Bhattacharya, A.; Gordon, D.F.; Maurer, R.A.
DNA 5, 299-304, 1986
A:Title: Analysis of the organization and nucleotide sequence of the chromosomal gene for the rat thyrotropin subunit mRNA levels in rat anterior pituitary gland.
A:Reference number: A90954; MUID:86300091
A:Accession: A90954
A:Molecule type: DNA
A:Residues: 1-138 <CR2>
R:Croyle, M.L.; Maurer, R.A.
DNA 3, 231-236, 1984
A:Title: Thyroid hormone decreases thyrotropin subunit mRNA levels in rat anterior pituitary gland.
A:Reference number: A90946; MUID:84260949
A:Accession: A90946
A:Molecule type: mRNA
A:Residues: 1-138 <CR2>
A:Cross-references: GB:X01454; NID:g57416; PIDN:CAA25684.1; PID:g57417
A:Note: 16-Glu was also found
R:Carr, F.E.; Need, L.R.; Chin, W.W.
J. Biol. Chem. 262, 981-987, 1987
A:Title: Isolation and characterization of the rat thyrotropin beta-subunit gene. Differences between the rat and human genes.
A:Reference number: A26706; MUID:87109273
A:Accession: A26706
A:Molecule type: DNA
A:Residues: 1-15, 'G', 17-138 <CR2>
A:Cross-references: GB:M14450; GB:J02648; NID:g207532; PIDN:AAA99238.1; PID:g207534
R:Chin, W.W.; Muccini, J.A.
Biochem. Biophys. Res. Commun. 128, 1152-1158, 1985
A:Title: Evidence for a single rat thyrotropin-beta-subunit gene: Thyroidectomy increases the expression of the gene.
A:Reference number: I52206; MUID:85225532
A:Accession: I52206
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-15, 'G', 17-138 <RES>
A:Cross-references: GB:M10502; NID:g207535; PIDN:AAA42301.1; PID:g207536
C:Genetics: 54/3
A:Introns: 54/3
C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: glycoprotein; hormone; pituitary
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-132/Product: thyrotropin beta chain #status predicted <MAT>
F:133-138/Domain: carboxyl-terminal propeptide #status predicted <CPT>

F:22-47,36-72,39-103,51-125,87-115,105-108/disulfide bonds: #status predicted
F:43/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 93.0%; Score 159; DB 1; Length 138;
Best Local Similarity 93.3%; Pred. No. 5.3e-15;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 vctyrdshlyrtveipgcpvavpysypva 30
|||||
Db 71 VCTYRDSHYRTVEIPGCPHVAVPYSYVPA 100

RESULT 2
A29479
Thyrotropin beta chain precursor - mouse
N:Alternate names: thyroid stimulating hormone beta-subunit
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 16-Jul-1999
C:Accession: A29479; A28478; I49544; I49545; I49546; I49547; I59443
R:Gordon, D.F.; Wood, W.M.; Ridgway, E.C.
DNA 7, 17-26, 1988
A:Title: Organization and nucleotide sequence of the gene encoding the beta-subunit of the thyroid stimulating hormone.
A:Reference number: A29479; MUID:88166350
A:Accession: A29479
A:Molecule type: DNA
A:Residues: 1-138 <GOR>
A:Cross-references: GB:M22740; GB:M19138; NID:g202185; PIDN:AAA40493.1; PID:g202187
R:Wolf, O.; Kourides, I.A.; Gurr, J.A.
J. Biol. Chem. 262, 16596-16603, 1987
A:Title: Expression of the gene for the beta subunit of mouse thyrotropin results in the synthesis of a protein with the same amino acid sequence as the beta subunit of mouse thyrotropin.
A:Reference number: A28478; MUID:88059045
A:Accession: A28478
A:Molecule type: DNA
A:Residues: 1-138 <WOL>
A:Cross-references: GB:M20537; GB:J03480; NID:g202192; PIDN:AAA40494.1; PID:g202194
R:Wood, W.M.; Gordon, D.F.; Ridgway, E.C.
Mol. Endocrinol. 1, 875-883, 1987
A:Title: Expression of the beta-subunit gene of Murine thyrotropin results in multiple forms of the protein.
A:Reference number: I49544; MUID:91042575
A:Accession: I49544
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-8 <RES>
A:Cross-references: GB:M35719; NID:g192215; PIDN:AAA37307.1; PID:g192216
A:Accession: I49545
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-8 <RE2>
A:Cross-references: GB:M35720; NID:g192217; PIDN:AAA37308.1; PID:g192218
A:Accession: I49546
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-8 <RE3>
A:Cross-references: GB:M35721; NID:g192219; PIDN:AAA37309.1; PID:g192220
A:Accession: I49547
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-8 <RE4>
A:Cross-references: GB:M35723; NID:g192221; PIDN:AAA37310.1; PID:g192222
R:Kourides, I.
Recent Prog. Horm. Res. 40, 79-120, 1984
A:Title: The regulation and organization of thyroid stimulating hormone genes.
A:Reference number: I59443; MUID:85015962
A:Accession: I59443
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-138 <RES>
A:Cross-references: GB:M54943; NID:g202182; PIDN:AAA40492.1; PID:g202183
C:Genetics: 54/3
A:Introns: 54/3
C:Superfamily: pituitary glycoprotein hormone beta chain

C;Keywords: glycoprotein; hormone; pituitary
F;1-20/Domain: signal sequence #status predicted <SIG>
F;1-138/Product: thyrotropin beta chain #status predicted <MAT>
F;22-47,36-72,39-103,51-125,87-115,105-108/Disulfide bonds: #status predicted

Query Match 91.2%; Score 156; DB 2; Length 138;
Best Local Similarity 90.0%; Pred. No. 1.4e-14;
Matches 27; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpplhvapysypva 30
|||||:|||||:|||||:|||||:|||||
Db 71 VCTYRDFMYKTAEIPGCPHVTPTFSYPVA 100

RESULT 3
TTPOB
Thyrotropin beta chain precursor - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 24-Apr-1984 #sequence_revision 23-Aug-1997 #text_change 31-Oct-1997
C;Accession: B30339; A01491
R;Hirai, T.; Takikawa, H.; Kato, Y.
Mol. Cell. Endocrinol. 63, 209-217, 1989
A;Title: Molecular cloning of cDNAs for precursors of porcine pituitary glycoprotein hormone
A;Reference number: A30339; MUID:89325834
A;Accession: B30339
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-138 <HI2>
R;Naghuin-Rogister, G.; Hennen, G.; Closset, J.; Kopeyan, C.
Eur. J. Biochem. 61, 157-163, 1976
A;Title: Porcine thyrotropin. The amino-acid sequence of the alpha and beta subunits.
A;Reference number: A01491; MUID:76092029
A;Accession: A01491
A;Molecule type: protein
A;Residues: 21-41, 'V', '43, 'S', '45-56, 'D', '58-108, 'D', '110-128, 'E', '130-132 <MAC>
A;Note: some ambiguity was found in the compositions of peptides for positions 39-54
C;Superfamily: pituitary glycoprotein hormone beta chain
C;Keywords: glycoprotein
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-132/Product: thyrotropin beta chain #status experimental <MAT>
F;133-138/Domain: carboxyl-terminal propeptide #status predicted <CPT>
F;22-47,36-72,39-103,51-125,87-115,105-108/Disulfide bonds: #status predicted
F;43/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 90.1%; Score 154; DB 1; Length 138;
Best Local Similarity 86.7%; Pred. No. 2.6e-14;
Matches 26; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpplhvapysypva 30
|||||:|||||:|||||:|||||:|||||
Db 71 VCTYRDFMYKTAEIPGCPHVTPTFSYPVA 100

RESULT 4
TTPOB
Thyrotropin beta chain precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Apr-1984 #sequence_revision 23-Aug-1997 #text_change 18-Jun-1999
C;Accession: I45985; A94673; S74109; A01490
R;Maurer, R.A.; Croyle, M.L.; Donelson, J.E.
J. Biol. Chem. 259, 5024-5027, 1984
A;Title: The sequence of a cloned cDNA for the beta subunit of bovine thyrotropin predicted
A;Reference number: I45985; MUID:84185607
A;Accession: I45985
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-138 <MAU>
A;Cross-references: GB:K01939; NID:g163792; PIDN:AAA30796.1; PID:g163793
R;Liao, T.H.; Pierce, J.G.
J. Biol. Chem. 246, 850-865, 1971
A;Title: The primary structure of bovine thyrotropin.

A;Reference number: A94673; MUID:71111428
A;Accession: A94673
A;Molecule type: protein
A;Residues: 21-54, 'B', '56, 'B', '58-108, 'B', '110, 'B', '112-113, 'B', '115-133 <LIA>
A;Note: two types of chains are found in pure preparations; the other type lacks the
R;Shome, B.; Liao, T.H.; Howard, S.M.; Pierce, J.G.
J. Biol. Chem. 246, 833-849, 1971
A;Title: The primary structure of bovine thyrotropin. I. Isolation and partial sequen
A;Reference number: A92098; MUID:71111427
A;Contents: annotation; partial sequence
R;Fairlie, W.D.; Stanton, P.G.; Hearn, M.T.W.
Eur. J. Biochem. 240, 622-627, 1996
A;Title: Contribution of specific disulfide bonds to two epitopes of thyrotropin beta
A;Reference number: S74109; MUID:97008953
A;Accession: S74109
A;Molecule type: protein
A;Residues: 21-27;35-38;39-50;51-54;65-74;81-89;95-104;105-107;108-121;125-130 <FAI>
C;Superfamily: pituitary glycoprotein hormone beta chain
C;Keywords: glycoprotein
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-133/Product: thyrotropin beta chain #status experimental <MAT>
F;134-138/Domain: carboxyl-terminal propeptide #status predicted <PRO>
F;22-72,36-87,39-125,47-103,51-105,108-115/Disulfide bonds: (version 2) #status exper
F;22-47,36-72,39-103,51-125,87-115,105-108/Disulfide bonds: (version 1) #status predi
F;43/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 88.3%; Score 151; DB 1; Length 138;
Best Local Similarity 83.3%; Pred. No. 6.8e-14;
Matches 25; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpplhvapysypva 30
|||||:|||||:|||||:|||||:|||||
Db 71 VCTYRDFMYKTAEIPGCPHVTPTFSYPVA 100

RESULT 5
S34148
Thyrotropin beta chain - European eel
C;Species: Anguilla anguilla (European eel)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Aug-1998
C;Accession: S34148
R;Salmon, C.; Marchelidon, J.; Fontaine, Y.A.; Huet, J.C.; Querat, B.
submitted to the EMBL data library, June 1993
A;Description: Cloning and sequence of thyrotropin (beta)-subunit of a teleost fish t
A;Reference number: S34148
A;Accession: S34148
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-147 <SAL>
A;Cross-references: EMBL:X73493
C;Superfamily: pituitary glycoprotein hormone beta chain

Query Match 64.3%; Score 110; DB 2; Length 147;
Best Local Similarity 65.5%; Pred. No. 3.4e-08;
Matches 19; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 cttyrdfiyrtveipgcpplhvapysypva 30
|||||:|||||:|||||:|||||:|||||
Db 72 CTYQAVEYRTAEIPGCPHLPFRFSYPVA 100

RESULT 6
A48194
Thyrotropin beta chain - rainbow trout
C;Species: Oncorhynchus mykiss (rainbow trout)
C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999
C;Accession: A48194
R;Ito, M.; Koide, Y.; Takamatsu, N.; Kawauchi, H.; Shiba, T.
Proc. Natl. Acad. Sci. U.S.A. 90, 6052-6055, 1993
A;Title: cDNA cloning of the beta subunit of teleost thyrotropin.
A;Reference number: A48194; MUID:93317622

A:Accession: S29677
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-144 <CH1>
A:Cross-references: EMBL:X59888; NID:462619; PIDN:CAA42542.1; PID:962620
R:Chang, Y.S.; Huang, C.J.; Huang, F.L.; Lo, T.B.
Int. J. Pept. Protein Res. 32, 556-564, 1988
A:Title: Primary structures of carp gonadotropin subunits deduced from cDNA
A:Reference number: JK0024; MUID:89233593
A:Accession: J00462
A:Molecule type: mRNA
A:Residues: 1-144 <CH2>
R:Chang, Y.S.; Huang, F.L.; Lo, T.B.

A;Status: preliminary

A:Molecule type: mRNA
A:Residues: 1-142 <SEK>
A:Cross-references: GB:M27154; NID:g213429; PIDN:AAA49409.1; PID:g213430
R:Itoh, H.; Suzuki, K.; Kawachi, H.
Gen. Comp. Endocrinol. 71, 438-451, 1988
A:Title: The complete amino acid sequences of beta-subunits of two distinct chum salmon
A:Reference number: S07216; MUID:89053031
A:Accession: S09344

A:Molecule type: protein
A:Residues: 24-72, I', 74-142 <ITO>
C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: glycoprotein; heterodimer; hormone; pituitary
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-142/Product: gonadotropin II beta chain #status experimental <WAT>
F:29-54,43-77,46-108,58-130,92-120,110-113/Disulfide bonds: #status predicted
F:33/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 57.9%; Score 99; DB 1; Length 142;
Best Local Similarity 56.7%; Pred. No. 1.1e-06;
Matches 17; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 1 vctyrdfyrtveipgcpplhvapysypva 30
||||| I I : | | | I I : | | |
Db 76 VCTYRDVRYETIRLPDCPPWDPHTVTPVA 105

RESULT 11
I50994
gonadotropin II beta subunit - Morone saxatilis
C:Species: Morone saxatilis
C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 16-Jul-1999
C:Accession: I50994
R:Hassin, S.; Elizur, A.; Zohar, Y.
J. Mol. Endocrinol. 15, 23-35, 1995
A:Title: Molecular cloning and sequence analysis of striped bass (Morone saxatilis) gonadotropin II beta subunit
A:Reference number: I50992; MUID:96020549
A:Accession: I50994
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-147 <RAS>
A:Cross-references: GB:L35096; NID:g522308; PIDN:RAAC38019.1; PID:g598255
C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match 57.9%; Score 99; DB 2; Length 147;
Best Local Similarity 60.0%; Pred. No. 1.1e-06;
Matches 18; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 1 vctyrdfyrtveipgcpplhvapysypva 30
||||| I I : | | | I I : | | |
Db 85 VCTYRDLHYKTFELPCPPGDPHTVTPVA 114

RESULT 12
I83048
FSH beta-subunit - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C:Accession: I83048
R:Kato, Y.; Ezashi, T.; Hirai, T.; Kato, T.
Zool. Sci. 7, 879-887, 1990
A:Title: Strain Difference in Nucleotide Sequences of Rat Glycoprotein Hormone Subunit c
A:Reference number: I60104
A:Accession: I83048
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-76 <RES>
A:Cross-references: GB:D00577; NID:g220739; PIDN:BA00455.1; PID:g220740
C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match 55.6%; Score 95; DB 2; Length 76;

Best Local Similarity 46.7%; Pred. No. 2.1e-06;
Matches 14; Conservative 7; Mismatches 9; Indels 0; Gaps 0;
OY 1 vctyrdfyrtveipgcpplhvapysypva 30
||||| I I : | | | I I : | | |
Db 15 VCTFRELVTYETIRLPGCARHSDSLTYTPVA 44

RESULT 13
UT0UB
lutropin beta chain precursor - human
N:Alternate names: interstitial cell-stimulating hormone (ICSH) beta chain; luteinizing hormone (LH) beta chain
C:Species: Homo sapiens (man)
C:Date: 13-Jul-1981 #sequence_revision 24-Nov-1999 #text_change 24-Nov-1999
C:Accession: I37994; I58013; A90604; A92759; A94466; A01497; B94552
R:Talmadge, K.; Vamvakopoulos, N.C.; Fiddes, J.C.
Nature 307, 37-40, 1984
A:Title: Evolution of the genes for the beta subunits of human chorionic gonadotropin
A:Reference number: I37231; MUID:84093590
A:Accession: I37994
A:Status: preliminary; translated from GB/EMBL/DBJ; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-141 <TAL>
A:Cross-references: GB:X00264; NID:g34351; PIDN:CAA25067.1; PID:e332193; PID:g2292893
R:Weiss, J.; Axelrod, L.; Whitcomb, R.W.; Harris, P.E.; Crowley, W.F.; Jameson, J.L.
N. Engl. J. Med. 326, 179-183, 1991
A:Title: Hypogonadism caused by a single amino acid substitution in the beta subunit
A:Reference number: I58013; MUID:92085985
A:Accession: I58013
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 72-73, R', 75-76 <WEI>
A:Cross-references: GB:S71273; NID:g240572; PIDN:AAD14960.1; PID:g4262812
A:Note: mutant sequence from patient with hypogonadism
R:Sairam, M.R.; Li, C.H.
Biochim. Biophys. Acta 412, 70-81, 1975
A:Title: Human pituitary lutropin. Isolation, properties, and the complete amino acid sequence
A:Reference number: A90604; MUID:76062547
A:Accession: A90604
A:Molecule type: protein
A:Residues: 21-38, 'Q', 40-75, 77-131, 133-134, 'H', 136-137 <SAI>
R:Shome, B.; Parlow, A.F.
J. Clin. Endocrinol. Metab. 36, 618-621, 1973
A:Title: The primary structure of the hormone-specific, beta subunit of human pituitary
A:Reference number: A92759; MUID:73090987
A:Accession: A92759
A:Molecule type: protein
A:Residues: 21-38, 'Q', 40-75, 77-131, 133-134, 'H', 136-137 <SHO>
R:Closset, J.; Hennen, G.; Lequin, R.M.
FEBS Lett. 29, 97-100, 1973
A:Title: Human luteinizing hormone the amino acid sequence of the beta subunit.
A:Reference number: A91389; MUID:73221227
A:Contents: annotation; partial sequence
R:Ward, D.N.
unpublished results, cited by Closset, J., Hennen, G., and Lequin, R.M., FEBS Lett. 2

A:Reference number: A94466
A:Accession: A94466
A:Molecule type: protein
A:Residues: 21-38, 'Q', 40-46 <WAR>
A:Note: 28-Val, 33-Arg, and 35-Thr were also found
R:Shome, B.; Parlow, A.F.
submitted to the Atlas, April 1975
A:Reference number: A94552
A:Contents: annotation; binding site
C:Genetics:
A:Gene: GDB:LHB
A:Cross-references: GDB:I19364; OMIM:152780
A:Map position: 19q13.3-19q13.3
C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: blocked amino end; glycoprotein; hormone; pituitary
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-141/Product: lutropin beta #status experimental <UTB>
F:21/Modified site: blocked amino end (Ser) (in mature form) (probably acetylated) #s

F:29-54,43-77,46-108,58-130,92-120,110-113/Disulfide bonds: #status predicted
F:50/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 55.6%; Score 95; DB 1; Length 141;
Best Local Similarity 53.3%; Pred. No. 3.9e-06;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgpcplhvpafsyypva 30
||||| : : : ||||| | : |||||
Db 76 VCTYRDVRFESIRLPGCPRGVDPVVSFPVA 105

RESULT 14

S74084
folliotropin beta chain - ostrich
C:Species: *Struthio camelus* (ostrich)
C>Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 17-Apr-1998
C:Accession: S74084
R:Koide, Y.; Papkoff, H.; Kawauchi, H.
Eur. J. Biochem. 240, 262-267, 1996
A:Title: Complete amino acid sequences of follitropin and lutropin in the ostrich, *Struthio camelus*
A:Reference number: S74084; MUID:97025333
A:Accession: S74084
A:Molecule type: protein
A:Residues: 1-106 <KOI>
A:Experimental source: pituitary glands
A:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: glycoprotein; heterodimer; hormone; pituitary
F:1-49,15-64,18-102,26-80,30-82,85-92/Disulfide bonds: #status predicted
F:5,22/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 55.0%; Score 94; DB 2; Length 106;
Best Local Similarity 51.7%; Pred. No. 4e-06;
Matches 15; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 2 ctgyrdfiyrtveipgpcplhvpafsyypva 30
||| : : ||||| | : |||||
Db 49 CTGKVVYETVKIPGRDHAEISLYPVA 77

RESULT 15

JC4526
folliotropin beta chain precursor - mouse
N:Alternate names: follicle-stimulating hormone beta chain
C:Species: *Mus musculus* (house mouse)
C>Date: 08-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 16-Jul-1999
C:Accession: JC4526
R:Kumar, T.R.; Kelly, M.; Mortrud, M.; Low, M.J.; Matzuk, M.M.
Gene 166, 333-334, 1995
A:Title: Cloning of the mouse gonadotropin beta-subunit-encoding genes, I. Structure of the mouse gonadotropin beta-subunit-encoding genes, I.
A:Reference number: JC4526; MUID:96125215
A:Accession: JC4526
A:Molecule type: mRNA
A:Residues: 1-130 <KUM>
A:Cross-references: GB:U12932; NID:G959515; PIDN:AAA92804.1; PID:G9595516
A:Experimental source: 129SvEv
C:Comment: This protein is a heterodimeric glycoprotein that consists of a common alpha subunit and an evolutionarily conserved beta subunit.
C:Genetics: 129SvEv
A:Gene: fsh beta

Query Match 54.4%; Score 93; DB 2; Length 130;
Best Local Similarity 46.7%; Pred. No. 6.8e-06;
Matches 14; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgpcplhvpafsyypva 30
||||| : : : ||||| | : |||||
Db 69 VCTFKELVYETVRLPGCARHSDSLYTPVA 98

RESULT 16

JC4527
lutinizing hormone beta chain precursor - mouse
C:Species: *Mus musculus* (house mouse)
C>Date: 15-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 16-Jul-1999
C:Accession: JC4527
R:Kumar, T.R.; Matzuk, M.M.
Gene 166, 335-336, 1995
A:Title: Cloning of the mouse gonadotropin beta-subunit-encoding genes, II. Structure of the mouse gonadotropin beta-subunit-encoding genes, II.
A:Reference number: JC4527; MUID:96125216
A:Accession: JC4527
A:Molecule type: mRNA
A:Residues: 1-141 <KUM>
A:Cross-references: GB:U25145; NID:G930344; PIDN:AAA92841.1; PID:G930345
A:Experimental source: 129SvEv
C:Comment: This protein is co-produced with follicle-stimulating hormone in pituitary gonadotrophs.
C:Genetics: 129SvEv
A:Gene: lh beta
A:Introns: 5/3; 61/3
A:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: hormone; reproduction
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-141/Product: lutinizing hormone beta chain #status predicted <MAT>

Query Match 54.4%; Score 93; DB 2; Length 141;
Best Local Similarity 53.3%; Pred. No. 7.3e-06;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgpcplhvpafsyypva 30
||||| : : : ||||| | : |||||
Db 76 VCTYRELAFASVRLPGCPGVDPIVSFPVA 105

RESULT 17

IS1232
gonadotropin II beta subunit - cherry salmon
C:Species: *Oncorhynchus masou* (cherry salmon)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: IS1232
R:Kato, Y.; Gen, K.; Maruyama, O.; Tomizawa, K.; Kato, T.
J. Mol. Endocrinol. 11, 275-282, 1993
A:Title: Molecular cloning of cDNAs encoding two gonadotropin beta subunits (GTH-I b and GTH-II b) from the cherry salmon, *Oncorhynchus masou*.
A:Reference number: IS1231; MUID:94197893
A:Accession: IS1232
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-142 <KAT>
A:Cross-references: GB:S69276; NID:G546263; PIDN:AAB30424.1; PID:G546264
C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match 54.4%; Score 93; DB 2; Length 142;
Best Local Similarity 53.3%; Pred. No. 7.4e-06;
Matches 16; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgpcplhvpafsyypva 30
||||| : : : ||||| | : |||||
Db 76 VCTYRDVRYETRLPGCPWDPVHTYTPVA 105

RESULT 18

FTROB
folliotropin beta chain - horse
C:Species: *Equus caballus* (domestic horse)
C>Date: 30-Jun-1979 #sequence_revision 30-Jun-1979 #text_change 08-Dec-1995
C:Accession: A01494

R:Fujiki, Y.; Rathnam, P.; Saxena, B.B.
J. Biol. Chem. 253, 5363-5368, 1978
A:Title: Amino acid sequence of the beta-subunit of the follicle-stimulating hormone from sheep
A:Reference number: A01494; MUID:78218213
A:Accession: A01494

A:Molecule type: protein
A:Residues: 1-118 <FUJ>
C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: glycoprotein
F:3-28,17-51,20-82,32-104,66-94,84-87/Disulfide bonds: #status predicted
F:7,24/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 53.8%; Score 92; DB 1; Length 118;
Best Local Similarity 44.8%; Pred. No. 8.5e-06;
Matches 13; Conservative 8; Mismatches 8; Indels 0; Gaps 0;
QY 2 ctvrdfiyrtveipgcpplhvayfsypva 30
||||: ||||:|||| | :|||||
Db 51 CTFKELVYETVKVPGCAHADSLSYTPVA 79

RESULT 19
FTSHB
N:Alternate names: follicle-stimulating hormone beta chain
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 18-Dec-1981 #sequence_revision 23-Aug-1996 #text_change 18-Jun-1999
A:Reference number: A04010; S05316; A01495
A:Accession: A04010; S05316; A01495
R:Guzman, K.; Miller, C.D.; Phillips, C.L.; Miller, W.L.
DNA Cell Biol. 10, 593-601, 1991

A:Title: The gene encoding ovine follicle-stimulating hormone beta: isolation, characterization and DNA sequence analysis of the cDNA for the precursor of ovine follicle-stimulating hormone beta
A:Reference number: A04010; MUID:92029621
A:Accession: A04010
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-129 <GUZ>
A:Cross-references: GB:S64745; NID:9238864; PIDN:AAB20317.1; PID:g238865
R:Mountford, P.S.; Bello, P.A.; Brandon, M.R.; Adams, T.E.
Nucleic Acids Res. 17, 6391, 1989
A:Title: Cloning and DNA sequence analysis of the cDNA for the precursor of ovine follicle-stimulating hormone beta
A:Reference number: S05316; MUID:89366671
A:Accession: S05316

A:Molecule type: mRNA
A:Residues: 1-129 <MOU>
R:Sairam, M.R.; Seidah, N.G.; Chretien, M.
Biochem. J. 197, 541-552, 1981
A:Title: Primary structure of the ovine pituitary follitropin beta-subunit.
A:Reference number: A01495; MUID:82113053
A:Accession: A01495
A:Molecule type: protein
A:Residues: 20-58,'B',60-63,'B',65-67,'T',69-106,'S',108-127,'ERZ' <SAI>
C:Superfamily: pituitary glycoprotein hormone beta chain
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-129/Product: follitropin beta chain #status predicted <MAT>
F:21-46,35-69,38-100,50-122,84-112,102-105/Disulfide bonds: #status predicted
F:25,42/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 53.8%; Score 92; DB 1; Length 129;
Best Local Similarity 44.8%; Pred. No. 9.2e-06;
Matches 13; Conservative 8; Mismatches 8; Indels 0; Gaps 0;
QY 2 ctvrdfiyrtveipgcpplhvayfsypva 30
||||: ||||:|||| | :|||||
Db 69 CTFKELVYETVKVPGCAHADSLSYTPVA 97

RESULT 20
FTPGB
follitropin beta chain precursor - pig

N:Alternate names: follicle-stimulating hormone beta chain
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 31-May-1979 #sequence_revision 23-Aug-1997 #text_change 18-Jun-1999
A:Reference number: A48169; I46582; A01496
R:Hirai, T.; Takikawa, H.; Kato, Y.
J. Mol. Endocrinol. 5, 147-158, 1990
A:Title: The gene for the beta subunit of porcine FSH: absence of consensus oestrogen binding site
A:Reference number: A48169; MUID:91063935
A:Accession: A48169
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-129 <HIR>
A:Cross-references: GB:D00621; NID:g217687; PIDN:BAA00499.1; PID:d1000954; PID:g21768

R:Kato, Y.
Mol. Cell. Endocrinol. 55, 107-112, 1988
A:Title: Cloning and DNA sequence analysis of the cDNA for the precursor of porcine FSH
A:Reference number: I46582; MUID:88196589
A:Accession: I46582
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 15-129 <KAT>
A:Cross-references: GB:M35676; NID:g164463; PIDN:AAA31039.1; PID:g164464
R:Closset, J.; Maghain-Rogister, G.; Hennen, G.; Strosberg, A.D.
Eur. J. Biochem. 86, 115-120, 1978

A:Title: Porcine follitropin. The amino-acid sequence of the beta subunit.
A:Reference number: A01496; MUID:78190610
A:Accession: A01496
A:Molecule type: protein
A:Residues: 21,'Z',23-30,'ZVKGLT',37-50,'T',52,'G',53,'B',55-58,'B',60-63,'B',65,'Z',67
C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: glycoprotein; heterodimer; hormone; pituitary
F:21-46,33-69,38-100,50-122,84-112,102-105/Disulfide bonds: #status predicted
F:25,42/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 53.8%; Score 92; DB 1; Length 129;
Best Local Similarity 44.8%; Pred. No. 9.2e-06;
Matches 13; Conservative 8; Mismatches 8; Indels 0; Gaps 0;
QY 2 ctvrdfiyrtveipgcpplhvayfsypva 30
||||: ||||:|||| | :|||||
Db 69 CTFKELVYETVKVPGCAHADSLSYTPVA 97

RESULT 21
A23550
follitropin beta chain precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 23-Aug-1987 #sequence_revision 23-Aug-1987 #text_change 16-Jul-1999
A:Reference number: A29816; A23550; A24914
R:Kim, K.E.; Gordon, D.F.; Maurer, R.A.
DNA 7, 227-233, 1988

A:Title: Nucleotide sequence of the bovine gene for follicle-stimulating hormone beta
A:Reference number: A29816; MUID:88283341
A:Accession: A29816
A:Molecule type: DNA
A:Residues: 1-129 <KIM>
A:Cross-references: GB:M83753; GB:M20185; NID:g163063; PIDN:AAA30528.1; PID:g163064
R:Zsch, F.S.; Mason, A.J.; Cooksey, K.; Mercado, M.; Shimasaki, S.
Proc. Natl. Acad. Sci. U.S.A. 83, 6618-6621, 1986
A:Title: Cloning and DNA sequence analysis of the cDNA for the precursor of the beta subunit of ovine follicle-stimulating hormone
A:Reference number: A23550; MUID:86313629
A:Accession: A23550
A:Molecule type: mRNA
A:Residues: 1-129 <ESC>

A:Cross-references: GB:M13383; NID:g163059; PIDN:AAA30526.1; PID:g163060
R:Maurer, R.A.; Beck, A.
DNA 5, 363-369, 1986
A:Title: Isolation and nucleotide sequence analysis of a cloned cDNA encoding the beta subunit of ovine follicle-stimulating hormone
A:Reference number: A24914; MUID:87053172
A:Accession: A24914
A:Molecule type: mRNA
A:Residues: 1-129 <MAU>

A:Cross-references: GB:M14853; NID:g163061; PIDN:AAA30527.1; PID:g163062
R:Watkins, P.C.; Eddy, R.; Beck, A.K.; Vellucci, V.; Leverone, B.; Tanzi, R.E.; Gusel
C:Superfamily: pituitary glycoprotein hormone beta chain
F:21-46,35-69,38-100,50-122,84-112,102-105/Disulfide bonds: #status predicted

Query Match 53.8%; Score 92; DB 2; Length 129;
Best Local Similarity 44.8%; Pred. No. 9.2e-06;
Matches 13; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

Qy 2 cttyrdfiyrtveipgcpbhvapyfsypva 30
||||: |||: ||| | : ||| |
Db 69 CTFKELVYETVVKPGCAHHDLSLTYPVA 97

RESULT 22
A32893
foliitropin beta chain precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 16-Jul-1999
C:Accession: A32893; A40060
R:Gharib, S.D.; Roy, A.; Wierman, M.E.; Chin, W.W.
DNA 8, 339-349, 1989
A:Title: Isolation and characterization of the gene encoding the beta-subunit of rat foli
A:Reference number: A32893; MUID:89356263
A:Accession: A32893
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-130 <GHA>
A:Cross-references: GB:M27044; GB:M27048; NID:g204179; PIDN:AAB60705.1; PID:g204181
R:Maurer, R.A.
Mol. Endocrinol. 1, 717-723, 1987
A:Title: Molecular cloning and nucleotide sequence analysis of complementary deoxyribonu
A:Reference number: A40060; MUID:91042555
A:Accession: A40060
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-130 <MAU>
A:Cross-references: GB:M36804
C:Superfamily: pituitary glycoprotein hormone beta chain
F:22-47,36-70,39-101,51-123,85-113,103-106/Disulfide bonds: #status predicted

Query Match 53.8%; Score 92; DB 2; Length 130;
Best Local Similarity 43.3%; Pred. No. 9.3e-06;
Matches 13; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcpbhvapyfsypva 30
||||: |||: ||| | : ||| |
Db 69 VCTFKELVYETVIRLPGCAHHDLSLTYPVA 98

RESULT 23
FTHUB
foliitropin beta chain precursor - human
N:Alternate names: follicle-stimulating hormone (FSH) beta chain
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1984 #sequence_revision 31-Dec-1993 #text_change 18-Jun-1999
R:Jameson, J.L.; Becker, C.B.; Lindell, C.M.; Habener, J.F.
Mol. Endocrinol. 2, 806-815, 1988
A:Title: Human follicle-stimulating hormone beta-subunit gene encodes multiple messenger
A:Reference number: A40920; MUID:89014411
A:Accession: A40920
A:Molecule type: DNA
A:Residues: 1-129 <JAM>
A:Cross-references: GB:M54912; GB:M54913; GB:M54914; NID:g182765; PIDN:AAB02868.1; PID:g
R:Keene, J.L.; Matzuk, M.M.; Ocan, T.; Fauser, B.C.J.M.; Galway, A.B.; Hsueh, A.J.W.; B
J. Biol. Chem. 264, 4769-4775, 1989
A:Title: Expression of biologically active human follitropin in Chinese hamster ovary ce
A:Reference number: A33266; MUID:89174746
A:Accession: A33266
A:Molecule type: DNA
A:Residues: 1-129 <KEE>

A:Cross-references: GB:M24538; GB:M24540; GB:J04612; NID:g341198; PIDN:AAAS2470.1; PI
R:Watkins, P.C.; Eddy, R.; Beck, A.K.; Vellucci, V.; Leverone, B.; Tanzi, R.E.; Gusel
DNA 6, 205-212, 1987
A:Title: DNA sequence and regional assignment of the human follicle-stimulating hormo
A:Reference number: A26959; MUID:87246070
A:Accession: A26959
A:Molecule type: DNA
A:Residues: 1-129 <WAT>
A:Cross-references: GB:M16647; NID:g182760; PIDN:AAAS2476.1; PID:g182762
R:Shome, B.; Parlow, A.F.; Liu, W.K.; Nahm, H.S.; Wen, T.; Ward, D.N.
J. Protein Chem. 7, 325-339, 1988
A:Title: A reevaluation of the amino acid sequence of human follitropin beta-subunit.
A:Reference number: A60892; MUID:89351581
A:Accession: A60892
A:Molecule type: protein
A:Residues: 19-129 <SHO>
R:Saxena, B.B.; Rathnam, P.
J. Biol. Chem. 251, 993-1005, 1976
A:Title: Amino acid sequence of the beta subunit of follicle-stimulating hormone from
A:Reference number: A92197; MUID:76120602
A:Accession: A92197
A:Molecule type: protein
A:Residues: 19-38; 'LV', '41-58', 'N', '60-128', 'QYPTALSY' <AUT>
R:Fujiki, Y.; Rathnam, P.; Saxena, B.B.
Biochim. Biophys. Acta 624, 428-435, 1980
A:Title: Studies on the disulfide bonds in human pituitary follicle-stimulating hormo
A:Reference number: A90632; MUID:81021713
A:Contents: annotation; disulfide bonds
R:Rathnam, P.; Tolvo, A.; Saxena, B.B.
Biochim. Biophys. Acta 708, 160-166, 1982
A:Title: Elucidation of the disulfide bond positions of the beta-subunit of human fol
A:Reference number: A21597; MUID:83075465
A:Contents: annotation; completion of disulfide bond assignments
R:Shome, B.; Parlow, A.F.
J. Clin. Endocrinol. Metab. 39, 203-205, 1974
A:Title: Human follicle stimulating hormone: first proposal for the amino acid sequen
A:Reference number: A92761; MUID:74262938
A:Contents: annotation; sequence
A:Note: this sequence differs substantially from that shown
C:Genetics:
A:Gene: GDB:FSHB
A:Cross-references: GDB:119955; OMIM:136530
A:Map position: 11p13-11p13
A:Introns: 53/3
C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: glycoprotein; heterodimer; hormone; pituitary
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-129/Product: follitropin beta chain #status experimental <MAT>
F:21-46,35-69,38-100,50-122,84-112,102-105/Disulfide bonds: #status experimental
F:25,42/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 53.2%; Score 91; DB 1; Length 129;
Best Local Similarity 44.8%; Pred. No. 1.3e-05;
Matches 13; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 2 cttyrdfiyrtveipgcpbhvapyfsypva 30
||||: |||: ||| | : ||| |
Db 69 CTFKELVYETVVKPGCAHHDLSLTYPVA 97

RESULT 24
I52320
testicular luteinizing hormone beta-subunit - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 16-Jul-1999
C:Accession: I52320
R:Zhang, F.P.; Rannikko, A.; Huhtaniemi, I.
Biochem. Biophys. Res. Commun. 210, 858-865, 1995
A:Title: Isolation and characterization of testis-specific cDNAs for luteinizing horm
A:Reference number: I52320; MUID:95283549
A:Accession: I52320
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1-139 <RES>
A:Cross-references: EMBL:U25653; NID:g915216; PIDN:AAC52249.1; PID:g915217
C:Genetics:
C:Gene: TLHB1
C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match 53.2%; Score 91; DB 2; Length 139;
Best Local Similarity 53.3%; Pred. No. 1.4e-05;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpvhvafsfypva 30
||||| : : : ||||| | : |||||
Db 74 VCTYRELRFASVRLPGCPGVDPIVSPFA 103

RESULT 25
UTRFB
N:Alternate names: interstitial cell-stimulating hormone (ICSH) beta chain; luteinizing
C:Species: Rattus norvegicus (Norway rat)
C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 24-Sep-1999
C:Accession: A01498; PQ0091; S42527; I83049
R:Jameson, L.; Chin, W.W.; Hollenberg, A.N.; Chang, A.S.; Habener, J.F.
J. Biol. Chem. 259, 15474-15480, 1984
A:Title: The gene encoding the beta-subunit of rat luteinizing hormone. Analysis of gene
A:Reference number: A01498; MUID:85080043
A:Accession: A01498
A:Molecule type: DNA
A:Residues: 1-141 <JAM>
A:Cross-references: EMBL:J00749; NID:g205175; PIDN:AAA96703.1; PID:g205176
R:Kato, Y.; Ezashi, T.; Hirai, T.; Kato, T.
submitted to JIPID, July 1990
A:Reference number: PQ0091
A:Accession: PQ0091
A:Molecule type: mRNA
A:Residues: 4-141 <KAT>
R:Chin, W.W.; Godine, J.E.; Klein, D.R.; Chang, A.S.; Tan, L.K.; Habener, J.F.
Proc. Natl. Acad. Sci. U.S.A. 80, 4649-4653, 1983
A:Title: Nucleotide sequence of the cDNA encoding the precursor of the beta subunit of
A:Reference number: S42527; MUID:83273673
A:Accession: S42527
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-141 <CHI>
A:Cross-references: GB:J00749; GB:M54957; NID:g205175; PIDN:AAA96703.1; PID:g205176
R:Kato, Y.; Ezashi, T.; Hirai, T.; Kato, T.
Zool. Sci. 7, 879-887, 1990
A:Title: Strain Difference in Nucleotide Sequences of Rat Glycoprotein Hormone Subunit c
A:Reference number: I60104
A:Accession: I83049
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 4-141 <RES>
A:Cross-references: GB:D00576; NID:g220807; PIDN:RAA00454.1; PID:g220808
C:Comment: The beta chain confers the specificity of the hormone.
C:Genetics:
A:Introns: 5/3; 61/3
C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: glycoprotein; hormone; pituitary
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-141/Product: lutropin beta #status predicted <LNB>
F:29-54,43-77,46-108,58-130,92-120,110-113/Disulfide bonds: #status predicted
F:33/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 53.2%; Score 91; DB 1; Length 141;
Best Local Similarity 53.3%; Pred. No. 1.4e-05;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpvhvafsfypva 30
||||| : : : ||||| | : |||||

Db 76 VCTYRELRFASVRLPGCPGVDPIVSPFA 105

RESULT 26

UTPGB
N:Alternate names: interstitial cell-stimulating hormone (ICSH) beta chain; luteinizing
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 24-Apr-1984 #sequence_revision 30-Jun-1993 #text_change 24-Nov-1999
C:Accession: A48170; A30322; A01501; A60584
R:Ezashi, T.; Hirai, T.; Kato, T.; Wakabayashi, K.; Kato, Y.
J. Mol. Endocrinol. 5, 137-146, 1990
A:Title: The gene for the beta subunit of porcine LH: clusters of GC boxes and CACCC
A:Reference number: A48170; MUID:91063934
A:Accession: A48170
A:Molecule type: DNA
A:Residues: 1-141 <EZA>
A:Cross-references: GB:D00579; NID:g217693; PIDN:BAA00457.1; PID:d1000912; PID:g21769
R:Kato, Y.; Hirai, T.
Mol. Cell. Endocrinol. 62, 47-53, 1989
A:Title: Cloning and DNA sequence analysis of the cDNA for the precursor of porcine LH
A:Reference number: A30322; MUID:89306142
A:Accession: A30322
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-141 <KAT>
R:Maghuln-Rogister, G.; Hennen, G.
Eur. J. Biochem. 39, 235-253, 1973
A:Title: Luteinizing hormone. The primary structures of the beta-subunit from bovine
A:Reference number: A91212; MUID:74075724
A:Accession: A01501
A:Molecule type: protein
A:Residues: 21-29, 31-39, 'D', 41-61, 'R', 63-82, 'I', 84-86, 'S', 88-121, 'PG', 124-133, 'P'
A:Note: 30-Arg was also found
A:Note: about half the chains lack one or both carboxyl-terminal leucines
R:Nomura, K.; Ohmura, K.; Nakamura, Y.; Horiba, N.; Shirakura, Y.; Sato, Y.; Ujihara,
Endocrinology 124, 712-719, 1989
A:Title: Porcine luteinizing hormone isoform(s): relationship between their molecular
A:Reference number: A60584; MUID:89107050
A:Accession: A60584
A:Molecule type: protein
A:Residues: 21-31; 137-139 <NOM>
A:Note: the lutropin beta chain is heterogeneous at the carboxyl end; this form lacks
C:Genetics:
A:Introns: 5/3; 61/3
C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: blocked amino end; glycoprotein; hormone
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-141/Product: lutropin beta chain #status experimental <MAT>
F:21/Modified site: blocked amino end (Ser) (in mature form) (probably acetylated) #s
F:29-54,43-77,46-108,58-130,92-120,110-113/Disulfide bonds: #status predicted
F:33/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 52.6%; Score 90; DB 1; Length 141;
Best Local Similarity 50.0%; Pred. No. 1.9e-05;
Matches 15; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpvhvafsfypva 30
||||| : : : ||||| | : |||||
Db 76 VCTYRELRFASVRLPGCPGVDPIVSPFA 105

RESULT 27

KTHOB
N:Alternate names: chorionic gonadotropin beta chain (CG); luteinizing hormone (LH) b
C:Species: Equus caballus (domestic horse)
C:Date: 14-Nov-1983 #sequence_revision 03-May-1996 #text_change 18-Jun-1999
C:Accession: A41917; A29304; A29305; A01503
R:Sherman, G.B.; Wolfe, M.W.; Farmerie, T.A.; Clay, C.M.; Threadgill, D.S.; Sharp, D.
Mol. Endocrinol. 6, 951-959, 1992
A:Title: A single gene encodes the beta-subunits of equine luteinizing hormone and ch


```

Best Local Similarity 50.0%; Pred. NO. 2.2e-05;
Matches 15; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcplhvpafysypva 30
||||| : : : ||||| | : ||||
Db 56 VCTYRZLRFASIRLPGCPGVDPVWSEFPVA 85

RESULT 29
A61465
lutropin beta chain - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 08-Dec-1995
C:Accession: A61465
R:Glenn, S.D.; Nahm, H.S.; Ward, D.N.
J. Protein Chem. 3, 259-273, 1984
A:Title: The amino acid sequence of the rabbit lutropin beta subunit.
A:Reference number: A61465
A:Accession: A61465
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-119 <GLE>
A:Note: the sequence from Fig. 1 is inconsistent with that from the abstract in havin
C:Superfamily: pituitary glycoprotein hormone beta chain
F:11-36,25-59,28-90,40-112,74-102,92-95/Disulfide bonds: #status predicted

Query Match 52.0%; Score 89; DB 2; Length 119;
Best Local Similarity 50.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcplhvpafysypva 30
||||| : : : ||||| | : ||||
Db 58 VCTYRELRFASIRLPGCPGVDPVWSEFPVA 87

RESULT 30
PN0141
lutropin beta chain - sperm whale
N:Alternate names: luteinizing hormone beta chain
C:Species: Physeter catodon (sperm whale)
C:Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 07-May-1999
C:Accession: PN0141
R:Pankov, Y.A.; Karasev, V.S.
Biokhimiia 49, 1004-1018, 1984
A:Title: Luteinizing hormone of the sperm-whale: amino acid sequence of reduced and
A:Reference number: PN0141; MUID:84281133
A:Accession: PN0141
A:Molecule type: protein
A:Residues: 1-118 <PAN>
A:Note: article in Russian with English abstract
C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: glycoprotein; hormone
F:9-34,23-57,26-88,38-110,72-100,90-93/Disulfide bonds: #status predicted
F:13/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 51.5%; Score 88; DB 2; Length 118;
Best Local Similarity 50.0%; Pred. No. 3e-05;
Matches 15; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcplhvpafysypva 30
||||| : : : ||||| | : ||||
Db 56 VCTYRZLRFASIRLPGCPGVDPVWSEFPVA 85

RESULT 31
S07092
gonadotropin beta chain - daggertooth pike conger
C:Species: Muraenesox cinereus (daggertooth pike conger)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S07092
R:Liu, C.S.; Huang, F.L.; Chang, Y.S.; Lo, T.B.

```


F:1-20/Domain: signal sequence #status predicted <SIG>
F:1-20/Product: lutropin beta #status experimental <LUT>
F:21-141/Modified site: blocked amino end (Ser) (in mature form) (probably acetylated) #status predicted
F:29-54,43-77,46-108,58-130,92-120,110-113/disulfide bonds: #status predicted
F:33/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 50.3%; Score 86; DB 1; Length 141;
Best Local Similarity 50.0%; Pred. No. 6.8e-05;

Matches 15; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpplhvapfysypva 30
||||| : : : : : ||||| | : |||||

Db 76 VCTYHELRFASVRLPGCPGVDPMVSPFVA 105

RESULT 36
UTSHB

lutropin beta chain precursor - sheep
N:Alternate names: interstitial cell-stimulating hormone (ICSH) beta chain; luteinizing
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 24-Apr-1984 #sequence_revision 02-May-1994 #text_change 24-Nov-1999
A:Accession: S09232; A02110; A90053; B61098; A01500
R:d'Angelo-Bernard, G.; Moumni, M.; Jutisz, M.; Counis, R.
Nucleic Acids Res. 18, 2175, 1990
A:Title: Cloning and sequence analysis of the cDNA for the precursor of the beta subunit
A:Reference number: S09232; MUID:90245669
A:Accession: S09232
A:Molecule type: mRNA
A:Residues: 1-141 <ANG>
A:Cross-references: EMBL:X52488; NID:g1319; PIDN:CAA36729.1; PID:g1320
R:Li, W.K.; Nahm, H.S.; Sweeney, C.M.; Holcomb, G.N.; Ward, D.N.
J. Biol. Chem. 247, 4365-4381, 1972
A:Title: The primary structure of ovine luteinizing hormone. II. The amino acid sequence
A:Reference number: A92110; MUID:72211145
A:Accession: A92110
A:Molecule type: protein
A:Residues: 21-58, 'P', '60-62', 'R', '64-121', 'PG', '124-125', 'E', '127-139 <LIU>
R:Sairam, M.R.; Sany, T.S.A.; Papkoff, H.; Li, C.H.
Arch. Biochem. Biophys. 153, 572-586, 1972
A:Title: The primary structure of ovine interstitial cell-stimulating hormone. II. The b
A:Reference number: A90053; MUID:73190035
A:Accession: A90053
A:Molecule type: protein
A:Residues: 21-29, 'E', '31-58', 'P', '60-62', 'R', '64-71', 'P', '72-80', 'O', '82-121', 'PG', '124-125', 'E', '12
R:Nomura, K.; Tsunasawa, S.; Ohmura, K.; Sakiyama, F.; Shizume, K.
Endocrinology 123, 700-712, 1988
A:Title: Renotropic activity in ovine luteinizing hormone isoform(s).
A:Reference number: A61098; MUID:88283534
A:Accession: B61098
A:Molecule type: protein
A:Residues: 21-39, 'N', '41-49', '64-78', 'V', '80-82', '84-106', '115-121', 'PG', '124-138 <NOM>
A:Note: this form was designated form beta-3; forms beta-1 and beta-2 each lack several
C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: blocked amino end; glycoprotein; hormone; pituitary
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-141/Product: lutropin beta chain #status experimental <MAT>
F:21/Modified site: blocked amino end (Ser) (in mature form) (partial) (probably acetyla
F:29-54, 43-77, 46-108, 58-130, 92-120, 110-113/Disulfide bonds: #status predicted
F:33/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 50.3%; Score 86; DB 1; Length 141;
Best Local Similarity 50.0%; Pred. No. 6.8e-05;
Matches 15; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpplhvapfysypva 30
||||| : : : : : ||||| | : |||||

Db 76 VCTYHELRFASVRLPGCPGVDPMVSPFVA 105

RESULT 37
I46949

lutinizing hormone beta-subunit - sheep
C:Species: Ovis sp. (sheep)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999
A:Accession: I46949
R:Brown, P.; McNeilly, J.R.; Wallace, R.M.; McNeilly, A.S.; Clark, A.J.
Mol. Cell. Endocrinol. 93, 157-165, 1993
A:Title: Characterization of the ovine LH beta-subunit gene: the promoter directs gonad
A:Reference number: I46949; MUID:93351742
A:Accession: I46949

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-141 <BRO>
A:Cross-references: GB:S64695; NID:g408240; PIDN:AAB27819.1; PID:g408241
C:Genetics:
A:Introns: 5/3; 61/3
C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match 50.3%; Score 86; DB 2; Length 141;
Best Local Similarity 50.0%; Pred. No. 6.8e-05;
Matches 15; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpplhvapfysypva 30
||||| : : : : : ||||| | : |||||

Db 76 VCTYHELRFASVRLPGCPGVDPMVSPFVA 105

RESULT 38
I37231

beta-gonadotropin - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-May-1997
A:Accession: I37231
R:Talmadge, K.; Vamvakopoulos, N.C.; Fiddes, J.C.
Nature 307, 37-40, 1984
A:Title: Evolution of the genes for the beta subunits of human chorionic gonadotropin
A:Reference number: I37231; MUID:84093590
A:Accession: I37231
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-145 <RES>
A:Cross-references: EMBL:X00266; NID:g29907
C:Genetics:
A:Introns: 41/3
C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match 50.3%; Score 86; DB 2; Length 145;
Best Local Similarity 50.0%; Pred. No. 7e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpplhvapfysypva 30
||||| : : : : : ||||| | : |||||

Db 56 VCNRYDRVFESIRLPGRGVNPPVSYAVA 85

RESULT 39
I37412

beta-gonadotropin - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-May-1997
A:Accession: I37412
R:Talmadge, K.; Vamvakopoulos, N.C.; Fiddes, J.C.
Nature 307, 37-40, 1984
A:Title: Evolution of the genes for the beta subunits of human chorionic gonadotropin
A:Reference number: I37231; MUID:84093590
A:Accession: I37412
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-145 <RES>
A:Cross-references: EMBL:X00265; NID:g31719
C:Genetics:
A:Introns: 41/3
C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match 50.3%; Score 86; DB 2; Length 145;
Best Local Similarity 50.0%; Pred. No. 7e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpplhvapfysypva 30
||||| : : : : : ||||| | : |||||

Db 56 VCNVDRVFESIRLPGCPGVNPVSYAVA 85

RESULT 40

KTHUB

Choriongonadotropin beta chain precursor - human

N;Alternate names: chorionic gonadotropin beta chain

C;Species: Homo sapiens (man)

C;Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 18-Jun-1999

C;Accession: A93230; I59972; I55224; I55250; I70007; I70008; A92303; A92181; A92142; PC1

R;Fiddes, J.C.; Goodman, H.M.

Nature 286, 684-687, 1980

A;Title: The cDNA for the beta-subunit of human chorionic gonadotropin suggests evolutio

A;Reference number: A93230; MUID:81012134

A;Accession: A93230

A;Molecule type: mRNA

A;Residues: 1-165 <PID>

A;Cross-references: GB:J00117; GB:M38559; GB:M54963; NID:g180436; PIDN:AAA96690.1; PID:9

R;Pollicastro, P.; Ovitt, C.E.; Hoshina, M.; Fukuoka, H.; Boothby, M.R.; Boime, I.

J. Biol. Chem. 258, 11492-11499, 1983

A;Title: The beta subunit of human chorionic gonadotropin is encoded by multiple genes.

A;Reference number: I59972

A;Accession: I69972

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-165 <POL>

A;Cross-references: GB:K03189; NID:g180450; PIDN:AAA53288.1; PID:g180453

A;Note: clone CG-beta-e

A;Accession: I55224

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-23, 'M', 25-136, 'A', 138-165 <PO>

A;Cross-references: GB:K03183; NID:g180442; PIDN:AAA53287.1; PID:g180444

A;Note: clone CG-beta-a

R;Pollicastro, P.F.; Daniels-McQueen, S.; Carle, G.; Boime, I.

J. Biol. Chem. 261, 5907-5916, 1986

A;Title: A map of the hCG beta-LH beta gene cluster.

A;Reference number: I55250; MUID:86195987

A;Accession: I55250

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-5 <PO>

A;Cross-references: GB:M13504; NID:g180419; PIDN:AAA52005.1; PID:g463088

A;Note: CG-beta-3 gene

A;Accession: I70007

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-5 <PO>

A;Cross-references: GB:M13505; NID:g180429; PIDN:AAA52008.1; PID:g463089

A;Note: CG-beta-6 gene

A;Accession: I70008

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-5 <RES>

A;Cross-references: GB:M13503; NID:g180432; PIDN:AAA52009.1; PID:g463090

A;Note: CG-beta-7 gene

R;Birken, S.; Fetherston, J.; Canfield, R.; Boime, I.

J. Biol. Chem. 256, 1816-1823, 1981

A;Title: The amino acid sequences of the prepeptides contained in the alpha and beta sub

A;Reference number: A92303; MUID:81117268

A;Accession: A92303

A;Molecule type: protein

A;Residues: 1-20 <BIR>

A;Note: The identity of the residue at position 19 could not be determined

R;Morgan, F.J.; Birken, S.; Canfield, R.E.

J. Biol. Chem. 250, 5247-5258, 1975

A;Title: The amino acid sequence of human chorionic gonadotropin. The alpha subunit and

A;Reference number: A92181; MUID:75211304

A;Accession: A92181

A;Molecule type: protein

A;Residues: 21-165 <MOR>

R;Carlsen, R.B.; Bahl, O.P.; Swaminathan, N.

J. Biol. Chem. 248, 6810-6827, 1973

A;Reference number: A92142; MUID:74011267

A;Accession: A92142

A;Molecule type: protein

A;Residues: 21-22, 'Q', 24-73, 'ZL', 76-140, 142-157, 'PB', 160-165, 'SLP' <CAR>

R;Shi, Z.P.; Du, G.G.; Li, W.X.; Li, S.Z.; Xu, Y.S.; Wang, Y.

Chinese Biochem. J. 6, 558-562, 1990

A;Title: The immunological characteristics of the enzymatic fragments of human chorio

A;Reference number: PC1016

A;Accession: PC1016

A;Molecule type: protein

A;Residues: 21-165 <SHI>

R;Birken, S.; Armstrong, E.G.; Kolk, M.A.G.; Cole, L.A.; Agosto, G.M.; Krichewsky, A

Endocrinology 123, 572-583, 1988

A;Title: Structure of the human chorionic gonadotropin beta-subunit fragment from pre

A;Reference number: A61097; MUID:88254680

A;Accession: A61097

A;Molecule type: protein

A;Residues: 26-32, 'X', 34-49, 'X', 51-60, 75-112 <BIR>

A;Note: this material from pregnancy urine lacks sialic acid in its carbohydrate and h

R;Kardana, A.; Bagshawe, K.D.; Coles, B.; Read, D.; Taylor, M.

Br. J. Cancer 67, 686-692, 1993

A;Title: Characterisation of UGP and its relationship with beta-core fragment.

A;Reference number: A56873; MUID:93229246

A;Accession: B56873

A;Molecule type: protein

A;Residues: 26-28, 'X', 30-32, 'X', 34-42, 'X', 44-45, 'X', 47-48, 75-76, 'X', 78-91, 'G', 93-102

A;Experimental source: urine

A;Note: sequence modified after extraction from NCBI backbone

A;Note: this material was designated urinary gonadotropin peptide (peak 2)

R;Lapthorn, A.J.; Harris, D.C.; Littlejohn, A.; Lustbader, J.W.; Canfield, R.E.; Mach

Nature 369, 455-461, 1994

A;Title: Crystal structure of human chorionic gonadotropin.

A;Reference number: A4674; MUID:94261179

A;Contents: annotation; X-ray crystallography, 3.0 angstroms; correction of disulfide

C;Genetics:

A;Gene: GDB:CGB

A;Cross-references: GDB:119055; OMIM:118860

A;Map position: 19q13.3-19q13.3

A;Introns: 5/3; 61/3

A;Note: the choriongonadotropin beta chain locus contains six genes (or pseudogenes)

C;Superfamily: pituitary glycoprotein hormone; pituitary

C;Keywords: glycoprotein; hormone; pituitary

F;1-20/Domain: signal sequence #status experimental <SIG>

F;21-165/Product: choriongonadotropin beta chain #status experimental <MAR>

F;29-77, 43-92, 46-130, 54-108, 58-110, 113-120/Disulfide bonds: #status experimental

F;33, 50/Binding site: carbohydrate (Asn) (covalent) #status experimental

F;138, 150/Binding site: carbohydrate (Ser) (covalent) #status predicted

F;141, 147, 152, 158/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 50.3%; Score 86; DB 1; Length 165;

Best Local Similarity 50.0%; Pred. No. 7.9e-05;

Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpvhvayfsypva 30

||| : : : ||| | | |

Db 76 VCNVDRVFESIRLPGCPGVNPVSYAVA 105

RESULT 41

S21196

lutropin beta chain - bullfrog

C;Species: Rana catesbeiana (bullfrog)

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997

C;Accession: S21196

R;Hayashi, H.; Hayashi, T.; Hanaoka, Y.

Eur. J. Biochem. 205, 105-110, 1992

A;Title: Amphibian lutropin from the bullfrog Rana catesbeiana. Complete amino acid s

A;Reference number: S21196; MUID:92209490

A;Accession: S21196

A;Status: preliminary

A;Molecule type: protein


```
A:Residues: 1-112 <HAY>
C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match      49.7%; Score 85; DB 2; Length 112;
Best Local Similarity 43.3%; Pred. No. 7.5e-05;
Matches 13; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgpcplhvapysfypva 30
Db 51 ICTYKEIRYDTIKLPDCLPGDPTFPVA 80

RESULT 42
S00512
lutropin beta chain precursor - dog (fragment)
N:Alternate names: lutinizing hormone beta chain
C:Species: Canis lupus familiaris (dog)
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 16-Jul-1999
C:Accession: S00512
R:Wolf, D.L.; Appleby, V.L.; Hjerrild, K.; Baker, A.R.; Talmadge, K.
Nucleic Acids Res. 15, 10602, 1987
A:Title: Nucleic acid and amino acid sequences of dog beta-LH: comparison to rat, cow and
A:Reference number: S00512; MUID:88096605
A:Accession: S00512
A:Molecule type: mRNA
A:Residues: 1-138 <WOL>
A:Cross-references: EMBL:Y00518; NID:9907; PIDN:CAA68572.1; PID:g860906
C:Superfamily: pituitary glycoprotein hormone beta chain
F:1-17/Domain: signal sequence (fragment) #status predicted <SIG>
F:18-138/Product: lutropin beta chain #status predicted <MAT>
F:26-51,40-74,43-105,55-127,89-117,107-110/Disulfide bonds: #status predicted

Query Match      49.7%; Score 85; DB 2; Length 138;
Best Local Similarity 46.7%; Pred. No. 9.2e-05;
Matches 14; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgpcplhvapysfypva 30
Db 73 VCTYHELHFASIRLPGCPGVPDMVSFPVA 102

RESULT 43
S00554
gonadotropin II beta subunit - mummichog
C:Species: Fundulus heteroclitus (mummichog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: I50554
R:Lin, Y.W.; Rupnow, B.A.; Price, D.A.; Greenberg, R.M.; Wallace, R.A.
Mol. Cell. Endocrinol. 85, 127-139, 1992
A:Title: Fundulus heteroclitus gonadotropins. 3. Cloning and sequencing of gonadotropic
A:Reference number: I50553; MUID:92405806
A:Accession: I50554
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-136 <LIN>
A:Cross-references: GB:M87015; NID:g213251; PIDN:AAB59963.1; PID:g213252
C:Genetics:
A:Gene: GTHIIbeta
C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match      48.5%; Score 83; DB 2; Length 136;
Best Local Similarity 53.3%; Pred. No. 0.00017;
Matches 16; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgpcplhvapysfypva 30
Db 74 VCTYGLYKTKTFEPCEPCGVPDPMVTPVA 103

RESULT 44
S07216
gonadotropin I beta chain - chum salmon
C:Species: Oncorhynchus keta (chum salmon)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S07216
R:Ittoh, H.; Suzuki, K.; Kawachi, H.
Gen. Comp. Endocrinol. 71, 438-451, 1988
A:Title: The complete amino acid sequences of beta-subunits of two distinct chum salmon
A:Reference number: S07216; MUID:89053031
A:Accession: S07216
A:Status: preliminary
```

```
KTAB
choriogonadotropin beta chain precursor - olive baboon
C:Species: Papio anubis, Papio hamadryas anubis (olive baboon)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 28-May-1999
C:Accession: A25808
R:Crawford, R.J.; Tregear, G.W.; Niall, H.D.
Gene 46, 161-169, 1986
A:Title: The nucleotide sequences of baboon chorionic gonadotropin beta-subunit genes
A:Reference number: A25808; MUID:87106851
A:Accession: A25808
A:Molecule type: mRNA
A:Residues: 1-165 <CRA>
A:Cross-references: GB:M14966; NID:q176572; PIDN:AAA35383.1; PID:q176573
C:Comment: There are at least five copies of CG-related genes and at least two of the
C:Superfamily: pituitary glycoprotein hormone beta chain
C:Keywords: glycoprotein; hormone; placenta; pregnancy maintenance
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-165/Product: chorogonadotropin beta chain #status predicted <CGB>
F:29-77,43-92,46-130,54-108,58-110,113-120/Disulfide bonds: #status predicted
F:33,50/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:140,147,152/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match      47.4%; Score 81; DB 1; Length 165;
Best Local Similarity 46.7%; Pred. No. 0.00039;
Matches 14; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgpcplhvapysfypva 30
Db 76 VCNREVREPSIRLPGCPGVPDMVSPVA 105

RESULT 45
I65235
testicular lutinizing hormone beta subunit - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 16-Jul-1999
C:Accession: I65235
R:Zhang, F.P.; Rannikko, A.; Huhtaniemi, I.
Biochem. Biophys. Res. Commun. 210, 858-865, 1995
A:Title: Isolation and characterization of testis-specific cDNAs for lutinizing horm
A:Reference number: I52320; MUID:95283549
A:Accession: I65235
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-80 <RES>
A:Cross-references: EMBL:U25803; NID:g904025; PIDN:AAC52251.1; PID:g904026
C:Genetics:
A:Gene: TLHB3
C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match      47.1%; Score 80.5; DB 2; Length 80;
Best Local Similarity 53.3%; Pred. No. 0.00023;
Matches 16; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

Qy 1 vctyrdfiyrtveipgpcplhvapysfypva 30
Db 16 VCTYRE-REASVRLPGCPGVPDIVSFPVA 44

RESULT 46
S07216
gonadotropin I beta chain - chum salmon
C:Species: Oncorhynchus keta (chum salmon)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S07216
R:Ittoh, H.; Suzuki, K.; Kawachi, H.
Gen. Comp. Endocrinol. 71, 438-451, 1988
A:Title: The complete amino acid sequences of beta-subunits of two distinct chum salmon
A:Reference number: S07216; MUID:89053031
A:Accession: S07216
A:Status: preliminary
```


RESULT 49
S34349
gonadotropin I beta chain - arctic cisco
C;Species: Coregonus autumnalis (arctic cisco)

Job time: 2250 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 14, 2000, 08:55:49 ; Search time 31.39 Seconds
(without alignments)
29.106 Million cell updates/sec

Title: BETA-CHAIN

Perfect score: 171

Sequence: 1 vctyrdlyrtveipgcplhvapyfsypva 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 83857 seqs, 30454973 residues

Total number of hits satisfying chosen parameters: 83856

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 99%

Listing first 1000 summaries

Database : SwissProt_38.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	159	93.0	101	1	TSHE_PHOSU
2	159	93.0	138	1	TSHE_RAT
3	156	91.2	138	1	TSHE_MOUSE
4	155	90.6	138	1	TSHE_CANFA
5	154	90.1	138	1	TSHE_LAMGL
6	154	90.1	138	1	TSHE_PIG
7	153	89.5	138	1	TSHE_HORSE
8	151	88.3	138	1	TSHE_BOVIN
9	149	86.6	134	1	TSHE_CHICK
10	143	86.1	147	1	TSHE_ANGAN
11	109	63.7	139	1	TSHE_SALSA
12	109	63.7	147	1	TSHE_ONCMY
13	107	62.6	115	1	GTH2_THUOB
14	101	59.1	141	1	GTH2_HYPMO
15	101	59.1	144	1	GTH2_CYPCA
16	101	59.1	146	1	GTH2_CTEID
17	99	57.9	142	1	GTH2_ONCKE
18	98	57.3	138	1	LSHB_MACRU
19	95	55.6	141	1	LSHB_HUMAN
20	94	55.0	106	1	FSHE_STRCA
21	93	54.4	130	1	FSHE_MOUSE
22	93	54.4	142	1	GTH2_ONCMA
23	92	53.8	118	1	FSHB_HORSE
24	92	53.8	129	1	FSHB_BOVIN
25	92	53.8	129	1	FSHE_PIG
26	92	53.8	129	1	FSHE_SHEEP
27	92	53.8	129	1	FSHE_TRIVU
28	92	53.8	130	1	FSHE_RAT
29	91	53.2	105	1	LSHB_MOUSE
30	91	53.2	129	1	FSHE_HUMAN
31	91	53.2	141	1	LSHB_RAT
32	90	52.6	138	1	GTHB_CLAGA
33	90	52.6	141	1	LSHB_PIG

34	52.6	169	1	LSHB_EQUBU
35	52.6	169	1	LSHB_HORSE
36	52.0	118	1	LSHB_BALAC
37	52.0	169	1	LSHB_EQUAS
38	51.5	118	1	LSHB_PHYSA
39	50.9	113	1	GTHB_MURCI
40	50.9	140	1	GTH2_ANGAN
41	50.9	142	1	GTH2_CORAU
42	50.9	142	1	GTHB_ONCTS
43	50.3	141	1	LSHB_BOVIN
44	50.3	141	1	LSHB_SHEEP
45	50.3	165	1	CGHB_HUMAN
46	49.7	112	1	LSHB_RANCA
47	49.7	138	1	LSHB_CANFA
48	48.5	136	1	GTH2_FUNHE
49	47.4	165	1	CGHB_PAPAN
50	40.1	137	1	GTH1_ONCKE
51	40.1	137	1	GTH1_ONCMA
52	39.8	164	1	CGHB_CALTJA
53	34.2	137	1	GTH1_CORAU
54	33.3	889	1	NODV_BRAJA
55	29.8	166	1	LSHB_COTJA
56	28.7	102	1	GTH1_THUOB
57	28.7	191	1	PGHD_FELCA
58	28.1	128	1	LSHB_STRCA
59	28.1	159	1	LSHB_MELGA
60	27.5	448	1	THDF_AQUAE
61	27.5	453	1	BIOA_AQUAE
62	27.5	494	1	GPPA_ECOLI
63	27.5	1670	1	CA34_HUMAN
64	27.2	446	1	ALLA_BACSU
65	26.9	366	1	THA_HUMAN
66	26.9	456	1	THDF_PSEPU
67	26.9	485	1	SYE_RHIME
68	26.6	114	1	GTH1_FUNHE
69	26.6	513	1	YD17_SCHPO
70	26.3	116	1	YK14_YEAST
71	26.3	139	1	RSFR_CHICK
72	26.3	348	1	E13B_PHAVU
73	26.3	416	1	PAX2_HUMAN
74	26.0	84	1	IAC2_HUMAN
75	26.0	486	1	RBL1_RHOSH
76	25.7	191	1	PGHD_BOVIN
77	25.7	256	1	HXDD_BRARE
78	25.7	308	1	GUTQ_ECOLI
79	25.7	323	1	PEX7_HUMAN
80	25.7	324	1	SG10_CAFEL
81	25.7	385	1	AAD_PHACH
82	25.7	454	1	THDF_ECOLI
83	25.7	461	1	THDF_HAEIN
84	25.7	1023	1	DPOL_ADEB3
85	25.4	576	1	YITO_YEAST
86	25.1	140	1	HVIG_MOUSE
87	25.1	282	1	SAST_VIBAN
88	25.1	287	1	PHNS_DESFR
89	25.1	318	1	PEX7_MOUSE
90	25.1	353	1	CCPA_ACEXY
91	25.1	399	1	IG1R_MOUSE
92	25.1	441	1	VG34_HSV11
93	25.1	443	1	PIRD_SCHPO
94	25.1	471	1	CA34_BOVIN
95	25.1	539	1	L114_CAEEL
96	25.1	618	1	PPCK_TREPA
97	25.1	653	1	EGLN_PIG
98	25.1	829	1	DPOL_BPT5
99	25.1	885	1	Y143_HUMAN
100	25.1	926	1	PERT_PIG
101	25.1	971	1	Y029_HUMAN
102	25.1	1299	1	VP75_HVSA
103	25.1	1616	1	APXL_HUMAN
104	24.9	369	1	TP6A_MEIJA
105	24.9	400	1	LIM3_MOUSE
106	24.9	851	1	OBP_HSV11

O46641	equus burch
P08751	equus cabal
P33088	balaenopter
P19794	equus asinu
P25330	physeter ca
P12837	muraesox
P27767	anguilla an
P48251	coregonus a
P07732	oncorhynch
P04651	bos taurus
P01231	ovis aries
P01233	homo sapien
P80071	rana catesb
P18842	canis famil
P30972	fundulus he
P07434	papio anubi
P48252	oncorhynch
P51500	callithrix
P48250	coregonus a
P15939	bradyrhizob
P45657	coturnix co
P37205	thunnus obe
P29487	felis silve
P80664	struthio ca
P45646	meleagris g
O67030	aquifex aeo
O65557	aquifex aeo
P25552	escherichia
O01955	homo sapien
O32137	bacillus su
P05111	homo sapien
P25755	pseudomonas
P15189	rhizobium m
P30971	fundulus he
O10240	schizosacch
P36078	saccharomyc
P30374	gallus gall
P23535	phaseolus v
O02962	homo sapien
P01155	homo sapien
P27997	rhodobacter
O02853	bos taurus
O40472	brachydanio
P17115	escherichia
O00628	homo sapien
P46568	caenorhabdi
O01752	phanerocha
P25522	escherichia
P43730	haemophilus
O72540	bovine aden
P40568	saccharomyc
P04219	mus muscu
P19829	vibrio angu
P18187	desulfovibr
P37697	acetobacter
O60751	mus muscu
O00155	ictaluriid h
P32747	schizosacch
Q28084	bos taurus
Q21446	caenorhabdi
O83159	tritonema p
P37176	sus scrofa
P19822	bacterioph
Q4156	homo sapien
P09933	sus scrofa
Q15032	homo sapien
P11282	herpesvirus
Q13796	homo sapien
O37815	methanococc
P30481	mus muscu
P10193	herpes simp

107	42.5	24.9	2241	1	TEGU_RCMVA	P16785 human cytom	180	40	23.4	750	1	PSM_HUMAN	Q04509 homo sapien
108	42.5	24.9	3027	1	POLG_PYFV1	Q05057 parsnip yel	181	40	23.4	852	1	RA54_SCHPO	P41410 schizosacch
109	42	24.6	55	1	IP1635 MONDELPHIS	P81635 monodelphis	182	40	23.4	1008	1	G110_YEAST	P36048 saccharomyc
110	42	24.6	191	1	PGHD_URSGAR	Q29562 ursus arcto	183	40	23.4	1123	1	VG39_HSV11	Q00143 ictalurid h
111	42	24.6	285	1	YHXC_BACSU	P40397 bacillus su	184	40	23.4	1168	1	DDX8_SCHPO	O42643 schizosacch
112	42	24.6	336	1	SG13_CABEL	P46567 caenorhabdi	185	40	23.4	1220	1	ATCP_PIG	P23220 sus scrofa
113	42	24.6	415	1	PAX2_MOUSE	P32114 mus musculu	186	40	23.4	1266	1	SVI_HUMAN	P41252 homo sapien
114	42	24.6	656	1	YK79_MYCTU	Q10687 mycobacteri	187	40	23.4	1312	1	PIPL_DROME	P25455 drosophila
115	42	24.6	680	1	GAG_SCVLA	P32503 saccharomyc	188	40	23.4	1584	1	YJ9G_YEAST	P41710 saccharomyc
116	42	24.6	2733	1	RRPB_CVMAS	P16342 murine coro	189	40	23.4	1663	1	CO3_RAT	P01026 rattus norv
117	41.5	24.3	285	1	DDH2_HUMAN	Q05865 homo sapien	190	39.5	23.1	106	1	HCAC_ECOLI	P77266 escherichia
118	41.5	24.3	325	1	YAL7_SCHPO	Q09929 schizosacch	191	39.5	23.1	195	1	YLM2_CAEEL	P34376 caenorhabdi
119	41.5	24.3	419	1	NEMO_HUMAN	Q096k9 homo sapien	192	39.5	23.1	265	1	CXB5_RAT	P36380 rattus norv
120	41.5	24.3	633	1	YLS3_CAEEL	P34388 caenorhabdi	193	39.5	23.1	266	1	CXB5_MOUSE	Q02738 mus musculu
121	41	24.0	116	1	GLHA_HUMAN	P01215 homo sapien	194	39.5	23.1	303	1	TCF1_MOUSE	Q00417 mus musculu
122	41	24.0	135	1	Y315_METJA	Q57763 methanococc	195	39.5	23.1	453	1	CARA_TRIVE	P87183 trichoderma
123	41	24.0	289	1	AMPR_RHOCA	P14145 rhodobacter	196	39.5	23.1	522	1	CP5B_CANTR	P30607 candida tro
124	41	24.0	310	1	EL13A_HORVU	P34742 hordeum vul	197	39.5	23.1	617	1	HEMA_MEAS1	P26028 measles vir
125	41	24.0	315	1	PHSS_DESBA	P13063 desulfovibr	198	39.5	23.1	620	1	HEMA_MEAS2	P28081 measles vir
126	41	24.0	328	1	P2Y6_RAT	Q63371 rattus norv	199	39.5	23.1	871	1	PC1_MOUSE	P06802 mus musculu
127	41	24.0	334	1	EL13B_HORVU	P15737 hordeum vul	200	39.5	23.1	1018	1	SVI_ARCFU	O29522 archaeoglob
128	41	24.0	335	1	EL13B_MAIZE	P49237 zea mays (m	201	39.5	23.1	1039	1	GUNB_CALSU	Q06706 saccharomyc
129	41	24.0	345	1	YJUN_ECOLI	P39400 escherichia	202	39.5	23.1	1349	1	IKI3_YEAST	P08706 trypanosoma
130	41	24.0	365	1	ERBS_SAGER	P14290 saccharopol	203	39.5	23.1	1530	1	RPC1_TRYBB	P13886 o'nyong-nyo
131	41	24.0	391	1	PAX2_BRARE	Q90268 brachydanio	204	39.5	23.1	2514	1	POLN_ONNVG	P06849 wheat dwarf
132	41	24.0	406	1	WCAL_SALTY	P26388 salmonella	205	39	22.8	90	1	Y10K_WDV	P23028 pseudonaja
133	41	24.0	480	1	DCD_RAT	P14173 rattus norv	206	39	22.8	133	1	PA2D_PSETE	P16159 trititum ae
134	41	24.0	506	1	SOID_NOCOP	Q04616 nocardia op	207	39	22.8	143	1	IA16_WHEAT	P33482 mus musculu
135	41	24.0	524	1	MYIN_SALTY	P37169 salmonella	208	39	22.8	146	1	PA2M_MOUSE	P23236 hordeum vul
136	41	24.0	656	1	RIP_MOUSE	Q60855 mus musculu	209	39	22.8	149	1	IAAB_HORVU	O66550 aquifex aeo
137	41	24.0	713	1	YHC4_YEAST	P38741 saccharomyc	210	39	22.8	160	1	HPK3_AQUAE	P51064 bartonella
138	41	24.0	782	1	OSTA_HAEIN	P44845 haemophilus	211	39	22.8	177	1	IPYR_BARBA	P77423 escherichia
139	41	24.0	1528	1	KEMI_YEAST	P22147 saccharomyc	212	39	22.8	181	1	HIFH_ECOLI	Q29095 sus scrofa
140	41	24.0	1964	1	NTC4_MOUSE	P13695 mus musculu	213	39	22.8	189	1	PGHD_PIG	P19562 bovine immu
141	41	24.0	2331	1	RRPL_MABVP	P31352 marburg vir	214	39	22.8	198	1	VIF_BIV06	P19563 bovine immu
142	41	24.0	2331	1	RRPL_MABVP	P35262 marburg vir	215	39	22.8	198	1	VIF_BIV27	P46908 bacillus su
143	40.5	23.7	154	1	Y4YB_RHISN	P55710 rhizobium s	216	39	22.8	238	1	FNR_BACSU	P73163 synechocyst
144	40.5	23.7	358	1	Y993_METJA	Q58400 methanococc	217	39	22.8	246	1	DLHH_SYNY3	P45363 escherichia
145	40.5	23.7	492	1	YGFB_ECOLI	P20403 escherichia	218	39	22.8	277	1	XAPA_ECOLI	P75044 mycoplasma
146	40.5	23.7	543	1	OPPA_ECOLI	P23843 escherichia	219	39	22.8	300	1	KPRS_MYCPN	O42068 gallus gall
147	40.5	23.7	816	1	P3K4_DICDI	P54676 dictyosteli	220	39	22.8	314	1	SIX3_CHICK	P42068 bacillus su
148	40.5	23.7	864	1	DNLI_CANAL	P52496 candida alb	221	39	22.8	319	1	HUTG_BACSU	P23813 mus musculu
149	40.5	23.7	873	1	PC1_HUMAN	P22413 homo sapien	222	39	22.8	323	1	HXB3_MOUSE	O95343 homo sapien
150	40.5	23.7	1107	1	POL2_RVVS	P36324 raspberry r	223	39	22.8	332	1	SIX3_MOUSE	Q62233 mus musculu
151	40.5	23.7	1391	1	RPC1_HUMAN	Q4802 homo sapien	224	39	22.8	333	1	SIX3_MOUSE	P03558 bacterioph
152	40.5	23.7	1555	1	GDE_RABIT	P35574 oryctolagus	225	39	22.8	365	1	VGL_BPIKE	P47011 saccharomyc
153	40	23.4	103	1	YBE8_YEAST	P38192 saccharomyc	226	39	22.8	380	1	GLG2_YEAST	P80011 rhizobium m
154	40	23.4	317	1	METR_ECOLI	P19797 escherichia	227	39	22.8	400	1	YAPQ_RHISN	P55620 rhizobium s
155	40	23.4	332	1	MDHC_ECHGR	Q04820 echinococcu	228	39	22.8	473	1	RBL1_THIPE	P28895 thiobacillu
156	40	23.4	347	1	HYPD_BRAJA	P31904 bradyrhizob	229	39	22.8	489	1	RBL_PORAE	Q09119 porphyridiu
157	40	23.4	371	1	NTF6_TOBAC	Q40531 nicotiana t	230	39	22.8	509	1	NFTL_RAT	P09414 rattus norv
158	40	23.4	379	1	HUPD_AZOCH	P42033 azotobacter	231	39	22.8	522	1	NFTA_CHICK	P17923 gallus gall
159	40	23.4	379	1	HYPD_ALCOV	P31903 alcaligenes	232	39	22.8	531	1	UD17_RAT	Q64633 rattus norv
160	40	23.4	379	1	HYPD_AZOVI	P1882 azotobacter	233	39	22.8	550	1	CRY5_DICDI	P21837 dictyosteli
161	40	23.4	385	1	HYPD_RHILV	P40598 rhizobium l	234	39	22.8	581	1	Y66K_CHLPS	O34023 chlamydia p
162	40	23.4	391	1	PAX5_HUMAN	Q02548 homo sapien	235	39	22.8	654	1	YED1_SCHPO	O13674 schizosacch
163	40	23.4	391	1	PAX5_MOUSE	Q02650 mus musculu	236	39	22.8	735	1	YMC0_YEAST	Q03722 saccharomyc
164	40	23.4	401	1	TAMI_POVHA	P30379 hamster pol	237	39	22.8	823	1	YNS2_CAEEL	P55862 xenopus lae
165	40	23.4	443	1	RBL_CALHE	Q05986 callitriche	238	39	22.8	833	1	YNS2_CAEEL	P34586 caenorhabdi
166	40	23.4	443	1	RBL_VERBO	P36490 verbenas bon	239	39	22.8	933	1	PERT_HUMAN	P07202 homo sapien
167	40	23.4	456	1	NIFK_METBA	P51754 methanosarc	240	39	22.8	1134	1	VGLM_HANTL	P16493 hantaan vir
168	40	23.4	465	1	YC14_KLEPN	Q48460 klebsiella	241	39	22.8	1135	1	VGLM_HANTL	P16853 hantaan vir
169	40	23.4	472	1	RBL1_HYDMR	Q59458 hydrogenovi	242	39	22.8	1135	1	VGLM_HANTL	P08668 hantaan vir
170	40	23.4	472	1	RBL_NITVU	Q59613 nitrobacter	243	39	22.8	1139	1	MA2X_HUMAN	P49541 homo sapien
171	40	23.4	473	1	RBL2_THIPE	Q07087 thiobacillu	244	39	22.8	1239	1	DPOG_HUMAN	P54098 homo sapien
172	40	23.4	476	1	RBL_MAIZE	P00874 zea mays (m	245	39	22.8	1332	1	XDH_HUMAN	P47389 homo sapien
173	40	23.4	488	1	RBL1_CYACA	P37393 cyanidium c	246	39	22.8	1339	1	DPOA_LEIDO	O00874 leishmania
174	40	23.4	511	1	MYIN_ECOLI	P75932 escherichia	247	39	22.8	1370	1	IGIR_RAT	P08120 drosophila
175	40	23.4	550	1	IDS_HUMAN	P22304 homo sapien	248	39	22.8	1775	1	CAL4_DROME	P08120 drosophila
176	40	23.4	631	1	PRKA_BACSU	P39134 bacillus su	249	39	22.8	2555	1	PPS3_BACSU	P3947 bacillus su
177	40	23.4	682	1	RPOC_ORYSA	P12092 oryza sativ	250	39	22.8	2647	1	ABP2_HUMAN	P21333 homo sapien
178	40	23.4	683	1	RPOC_MAIZE	P16024 zea mays (m	251	39	22.8	3712	1	LMA_DROME	Q00174 drosophila
179	40	23.4	690	1	HRPO_BURSO	P35656 burkholderi	252	39	22.8				

253	38.5	22.5	146	1	PA2M_RAT	P14423 rattus norv	326	38	22.2	756	1	AOC2_HUMAN	O75106 homo sapien
254	38.5	22.5	253	1	SOJ_BACSU	P37522 bacillus su	327	38	22.2	769	1	ITB2_HUMAN	P05107 homo sapien
255	38.5	22.5	259	1	SG11_CAEL	P45569 caenorhabdi	328	38	22.2	819	1	FGR1_CHICK	P21804 gallus gall
256	38.5	22.5	281	1	YBT1_CAEL	P18161 caenorhabdi	329	38	22.2	821	1	CAN3_HUMAN	P20807 homo sapien
257	38.5	22.5	283	1	PHAB_PSEOL	P26495 pseudomonas	330	38	22.2	823	1	YRR2_CAEL	Q09345 caenorhabdi
258	38.5	22.5	453	1	CARA_NEUCR	P22572 neurospora	331	38	22.2	892	1	NIRA_EMENI	P28348 emerichella
259	38.5	22.5	493	1	TRPE_PSEPU	P20579 pseudomonas	332	38	22.2	909	1	CTIA_FUSSO	P52958 fusarium so
260	38.5	22.5	500	1	XYNB_THESJ	O30360 thermoanaer	333	38	22.2	1040	1	YO43_CAEL	P34681 caenorhabdi
261	38.5	22.5	504	1	XYNB_BACST	Q92fm2 bacillus st	334	38	22.2	1051	1	ULK1_MOUSE	O70405 mus musculus
262	38.5	22.5	506	1	DHA4_YEAST	P54114 saccharomyc	335	38	22.2	1063	1	YNH6_CAEL	P32744 caenorhabdi
263	38.5	22.5	506	1	DHA5_YEAST	P47771 saccharomyc	336	38	22.2	1254	1	PHOG_YEAST	P15801 saccharomyc
264	38.5	22.5	541	1	OPPA_HAEIN	P71370 haemophilus	337	38	22.2	1263	1	CA24_ASCSU	P27393 ascaris suu
265	38.5	22.5	717	1	VGLH_HSVSA	P16492 herpesvirus	338	38	22.2	1263	1	SUIS_HUMAN	P14410 homo sapien
266	38.5	22.5	743	1	BGAL_THEET	P77989 thermoanaer	339	38	22.2	1839	1	ANXC_HUMAN	Q01485 homo sapien
267	38.5	22.5	884	1	TERT_YEAST	Q06163 saccharomyc	340	38	22.2	1856	1	MGA_HUMAN	Q43451 homo sapien
268	38.5	22.5	896	1	DSC3_HUMAN	P14574 homo sapien	341	38	22.2	2164	1	POLG_HRV89	P07210 human rhino
269	38.5	22.5	1216	1	AEQP_RAT	Q63191 rattus norv	342	38	22.2	3924	1	ANKB_HUMAN	Q01484 homo sapien
270	38.5	22.5	1608	1	RRPL_BDV	P25639 borina disea	343	37.5	21.9	150	1	VGD_BPAL3	P08765 bacterioph
271	38.5	22.5	1758	1	CA24_CAEL	P17140 caenorhabdi	344	37.5	21.9	150	1	VGD_BPPHX	Q38039 bacterioph
272	38.5	22.5	2491	1	MPRI_HUMAN	P17177 homo sapien	345	37.5	21.9	151	1	VGD_BPPHX	Q30637 bacterioph
273	38.5	22.5	4544	1	LRP1_HUMAN	Q07954 homo sapien	346	37.5	21.9	152	1	VGD_BPG4	P03638 bacterioph
274	38	22.2	112	1	VP26_HSV11	P10219 herpes simp	347	37.5	21.9	168	1	IAO3_WHEAT	P17314 triticum ae
275	38	22.2	155	1	ITRF_MAIZE	P01088 zea mays (m	348	37.5	21.9	174	1	CU22_BOMMO	O02388 bombyx mori
276	38	22.2	175	1	COAG_CARRO	P03997 carnoscor	349	37.5	21.9	232	1	PFS_ECOLI	P24247 escherichia
277	38	22.2	181	1	CASK_MOUSE	P06796 mus musculu	350	37.5	21.9	238	1	VGO2_VZVD	P09267 varicella-z
278	38	22.2	186	1	PTH_HELPY	P56077 helicobacte	351	37.5	21.9	298	1	DAPA_HAEIN	P43797 haemophilus
279	38	22.2	207	1	YADS_ECOLI	P37027 escherichia	352	37.5	21.9	342	1	ISIA_SYNP7	P15347 synechococc
280	38	22.2	210	1	YACE_NEIGO	Q50962 neisseria g	353	37.5	21.9	351	1	ARGC_SYNY3	P54899 synechocyst
281	38	22.2	221	1	NUOB_ECOLI	P33598 escherichia	354	37.5	21.9	353	1	YBGO_ECOLI	P75748 escherichia
282	38	22.2	220	1	TEF_PROSU	P97516 phodopus su	355	37.5	21.9	380	1	FOS_MOUSE	P01101 mus musculu
283	38	22.2	225	1	CRP_CAVPO	P49254 cavia porce	356	37.5	21.9	387	1	FD61_SOYBN	P48630 glycine max
284	38	22.2	225	1	LAPB_PASHA	P32181 pasteurellia	357	37.5	21.9	405	1	CGL_HUMAN	P32929 homo sapien
285	38	22.2	244	1	YRT2_CAEL	Q10045 caenorhabdi	358	37.5	21.9	468	1	COGI_RABIT	P13943 oryctolagus
286	38	22.2	261	1	TEF_RAT	P41224 rattus norv	359	37.5	21.9	506	1	AMT2_ECOLI	P26612 escherichia
287	38	22.2	292	1	GLMT_RAT	P13255 rattus norv	360	37.5	21.9	519	1	AMID_SYNY3	Q55424 synechocyst
288	38	22.2	297	1	KPBS_MYCGE	P47304 mycoplasma	361	37.5	21.9	529	1	Y4JF_RHISN	P55506 rhizobium s
289	38	22.2	316	1	E13E_HORVU	Q02438 hordeum vul	362	37.5	21.9	529	1	Y243_ARCFU	O29996 archaeoglob
290	38	22.2	325	1	Y272_SYNY3	P73893 synechocyst	363	37.5	21.9	558	1	Y653_HUMAN	O75144 homo sapien
291	38	22.2	333	1	REFB_SALTY	P26401 salmonella	364	37.5	21.9	566	1	DCPY_EMENI	Q62931 rattus norv
292	38	22.2	350	1	YJFQ_ECOLI	P39286 escherichia	365	37.5	21.9	566	1	DPY4_MOUSE	P87208 emerichella
293	38	22.2	364	1	A2HS_SHEEP	P29701 ovis aries	366	37.5	21.9	572	1	DPY4_MOUSE	O14531 homo sapien
294	38	22.2	383	1	DTC_HUMAN	P01880 homo sapien	367	37.5	21.9	572	1	CDRL_SCHPO	Q35098 mus musculu
295	38	22.2	394	1	MALA_BACST	Q45632 bacillus st	368	37.5	21.9	593	1	TRFE_RABIT	P07334 schizosacch
296	38	22.2	441	1	MALH_FUSMR	O06901 fusobacteri	369	37.5	21.9	695	1	SG2N_HUMAN	P19334 oryctolagus
297	38	22.2	458	1	TRKA_ECOLI	P23868 escherichia	370	37.5	21.9	713	1	YB9T_YEAST	Q13033 homo sapien
298	38	22.2	458	1	TRKA_SALTY	P39445 salmonella	371	37.5	21.9	807	1	YB9T_YEAST	P38148 saccharomyc
299	38	22.2	460	1	EFIA_AJECA	P40911 ajellomyces	372	37.5	21.9	888	1	YB9T_YEAST	Q29512 oryctolagus
300	38	22.2	468	1	PPAD_YEAST	P52290 saccharomyc	373	37.5	21.9	970	1	DIS3_SCHPO	Q29512 oryctolagus
301	38	22.2	471	1	RBL2_CHRVI	P22859 chromatium	374	37.5	21.9	1032	1	VG07_BFT4	P37202 schizosacch
302	38	22.2	471	1	RBL2_CHRVI	P22859 chromatium	375	37.5	21.9	1032	1	VG07_BFT4	P19081 bacterioph
303	38	22.2	480	1	DCD_MOUSE	P96486 synechococc	376	37.5	21.9	1081	1	PDR6_YEAST	P32767 saccharomyc
304	38	22.2	480	1	DCD_MOUSE	P22781 cavia porce	377	37.5	21.9	1091	1	JSN1_YEAST	P32767 saccharomyc
305	38	22.2	484	1	YLS5_CAEL	Q88533 mus musculu	378	37.5	21.9	1224	1	COPA_BOVIN	Q27954 bos taurus
306	38	22.2	520	1	UD17_MOUSE	P34390 caenorhabdi	379	37.5	21.9	1231	1	COPA_HUMAN	P53621 homo sapien
307	38	22.2	527	1	VL2_HPV38	Q62452 mus musculu	380	37.5	21.9	1231	1	GDE_HUMAN	P08603 homo sapien
308	38	22.2	530	1	P2B_EMENI	Q80912 human papil	381	37.5	21.9	1515	1	MPRI_MOUSE	Q07113 mus musculu
309	38	22.2	530	1	UD12_HUMAN	P36509 homo sapien	382	37	21.6	120	1	YQBH_BACSU	P45924 bacillus su
310	38	22.2	538	1	COX1_RHILE	Q08955 rhizobium l	383	37	21.6	122	1	IAAT_ELECO	P01087 eleusine co
311	38	22.2	543	1	PRAA_STRCO	P54739 streptomyce	384	37	21.6	132	1	Y16K_SSV1	P20195 sulfobolus
312	38	22.2	544	1	YB73_YEAST	P38319 saccharomyc	385	37	21.6	149	1	ETI1_CAVPO	P97740 cavia porce
313	38	22.2	556	1	TYRO_AGABI	Q42713 agaricus bi	386	37	21.6	162	1	HSBX_RAT	P97541 rattus norv
314	38	22.2	565	1	VGLF_SEN5	P27564 sendai viru	387	37	21.6	171	1	YVY6_CAEL	Q19005 caenorhabdi
315	38	22.2	565	1	VGLF_SENDF	P12575 sendai viru	388	37	21.6	180	1	CAS2_RABIT	P50418 oryctolagus
316	38	22.2	565	1	VGLF_SENDF	P04856 sendai viru	389	37	21.6	182	1	IDI_ECOLI	Q46822 escherichia
317	38	22.2	565	1	VGLF_SENDF	P04856 sendai viru	390	37	21.6	194	1	YMFQ_ECOLI	P75982 escherichia
318	38	22.2	565	1	VGLF_SENDF	P04856 sendai viru	391	37	21.6	209	1	GTP_BOVIN	P28801 bos taurus
319	38	22.2	572	1	DPV1_MOUSE	P97427 mus musculu	392	37	21.6	222	1	HUTG_KLEAE	P19452 klebsiella
320	38	22.2	572	1	DPV1_MOUSE	Q62950 rattus norv	393	37	21.6	224	1	YMY5_YEAST	Q03144 saccharomyc
321	38	22.2	584	1	YMB3_YEAST	Q04228 saccharomyc	394	37	21.6	247	1	TRV4_RAT	P12788 rattus norv
322	38	22.2	598	1	TFDB_ALCEU	P27138 alcaligenes	395	37	21.6	255	1	E13B_SOYBN	P52395 glycine max
323	38	22.2	638	1	LTK2_HUMAN	P53671 homo sapien	396	37	21.6	259	1	HXC9_FUGRU	O42502 fugu rubrip
324	38	22.2	650	1	SCAA_BOVIN	P55270 bos taurus	397	37	21.6	260	1	FLGG_ECOLI	P75939 escherichia
325	38	22.2	656	1	UL25_HCMVA	P16761 human cytom	398	37	21.6	260	1	FLAG_SALTY	P16439 salmonella

399	37	21.6	261	1	TEF_HUMAN	Q10587	homo sapien	472	37	21.6	1155	1	IF2P_METJA	Q57710	methanococc
400	37	21.6	303	1	ATN6_BUFMA	P30715	bufo marinu	473	37	21.6	1238	1	DPOG_MOUSE	P54099	mus musculus
401	37	21.6	306	1	KHSE_SYNY3	P73646	synechocyst	474	37	21.6	1256	1	FINC_CHICK	P11722	gallus gall
402	37	21.6	310	1	DCHS_LACS3	P00862	lactobacilli	475	37	21.6	1319	1	SSM4_YEAST	P40318	saccharomyc
403	37	21.6	317	1	METR_SALTY	P05984	salmonella	476	37	21.6	1459	1	YFIM_CABEL	P21874	caenorhabdi
404	37	21.6	335	1	YQ25_YEAST	P53238	saccharomyc	477	37	21.6	1687	1	P2142_HUMAN	P52746	homo sapien
405	37	21.6	357	1	FL3H_MATIN	Q05965	mathioliola i	478	37	21.6	1914	1	STCK_EMENI	Q00706	emericeella
406	37	21.6	358	1	MBHS_AZQVI	P21950	azotobacter	479	37	21.6	4303	1	PKDL_HUMAN	P98161	homo sapien
407	37	21.6	359	1	ACOX_ALCEU	P27748	alcaligenes	480	36.5	21.3	110	1	INS_CANFA	P01321	canis famli
408	37	21.6	365	1	VGL_EPT22	P15418	bacterioph	481	36.5	21.3	145	1	IAO1_WHEAT	P16850	tritcum ae
409	37	21.6	368	1	GLN2_MAIZE	P38560	zea mays (m	482	36.5	21.3	145	1	IAAA_HORVU	P28041	hordeum vul
410	37	21.6	372	1	MBHT_ECOLI	Q46847	escherichia	483	36.5	21.3	175	1	IPYR_ECOLI	P17288	escherichia
411	37	21.6	373	1	HYPD_ECOLI	P24192	escherichia	484	36.5	21.3	180	1	UCRI_CHLIT	Q46136	chlorobium
412	37	21.6	375	1	MBHS_CITFR	Q46045	citrobacter	485	36.5	21.3	211	1	RGML_YEAST	Q00453	saccharomyc
413	37	21.6	377	1	FL3H_HORVU	P28038	hordeum vul	486	36.5	21.3	237	1	NDF3_HUMAN	Q02886	homo sapien
414	37	21.6	377	1	HYPD_RROCA	P26411	rhodobacter	487	36.5	21.3	270	1	AAKB_RAT	P80386	rattus norv
415	37	21.6	378	1	HRCA_SYNY3	P72795	synechocyst	488	36.5	21.3	291	1	C61A_MOUSE	P46737	mus musculus
416	37	21.6	402	1	CPXF_STRGO	P18327	streptomyce	489	36.5	21.3	302	1	YDAK_ECOLI	P77744	escherichia
417	37	21.6	402	1	DFP_SYNY3	P73881	synechocyst	490	36.5	21.3	321	1	AOC3_RAT	O08590	rattus norv
418	37	21.6	425	1	PURA_FUSNU	O68581	fusobacteri	491	36.5	21.3	348	1	MCR3_ECOLI	P15006	escherichia
419	37	21.6	436	1	PUR6_CHICK	P38024	gallus gall	492	36.5	21.3	351	1	SPIN_HABPV	Q05894	heliothis a
420	37	21.6	445	1	YM85_MCTU	Q50680	mycobacteri	493	36.5	21.3	387	1	STCV_EMENI	Q00727	emericeella
421	37	21.6	459	1	CHLN_PINCO	P26180	pinus conto	494	36.5	21.3	392	1	P2X3_HUMAN	P56373	homo sapien
422	37	21.6	461	1	EF12_XENLA	P17507	xenopus lae	495	36.5	21.3	397	1	P2X3_RAT	P49654	rattus norv
423	37	21.6	461	1	EF13_XENLA	P17508	xenopus lae	496	36.5	21.3	407	1	CHLP_SYNY3	Q55087	synechocyst
424	37	21.6	462	1	EF10_XENLA	P13549	xenopus lae	497	36.5	21.3	411	1	CAPA_BACAN	P19579	bacillus an
425	37	21.6	462	1	EF1A_BRARE	Q92005	brachydanio	498	36.5	21.3	437	1	WTML_YEAST	Q12363	saccharomyc
426	37	21.6	462	1	EF1A_CHICK	Q90835	gallus gall	499	36.5	21.3	446	1	Y902_MCTU	Q10360	mycobacteri
427	37	21.6	465	1	VP19_HSVB2	P28935	equine herp	500	36.5	21.3	450	1	A2AA_RAT	P22909	rattus norv
428	37	21.6	467	1	CHLN_PINTH	P41646	pinus thunb	501	36.5	21.3	460	1	FDFT_SCHPO	P36596	schizosacch
429	37	21.6	473	1	MPI3_HUMAN	P30307	homo sapien	502	36.5	21.3	463	1	YRB3_CABEL	Q09400	caenorhabdi
430	37	21.6	473	1	XYLA_CLOSR	P48790	clostridium	503	36.5	21.3	486	1	GTFA_STRMU	P10249	streptococc
431	37	21.6	482	1	CATL_ARATH	Q96528	arabidopsis	504	36.5	21.3	488	1	RBL_GUTHI	P14957	guillardia
432	37	21.6	496	1	C7B1_THLAR	P49264	thlaspi arv	505	36.5	21.3	506	1	VP6_RDV	P22473	rice dwarf
433	37	21.6	500	1	CBP3_ORYSA	P37891	oryza sativ	506	36.5	21.3	507	1	UL17_EBV	P20222	epstein-bar
434	37	21.6	502	1	MPI3_PIG	Q29029	sus scrofa	507	36.5	21.3	508	1	FAS3_DROME	P152278	drosophila
435	37	21.6	510	1	GLPD_PSEAE	P52111	pseudomonas	508	36.5	21.3	520	1	TR14_FUSSP	Q12612	fusarium sp
436	37	21.6	518	1	GIT1_YEAST	P25346	saccharomyc	509	36.5	21.3	532	1	YC18_HASIN	Q57251	haemophilus
437	37	21.6	555	1	WEEL_XENLA	P47817	xenopus lae	510	36.5	21.3	549	1	TF65_MOUSE	Q04207	mus musculu
438	37	21.6	556	1	CD19_HUMAN	P15391	homo sapien	511	36.5	21.3	570	1	YGR0_YEAST	P32109	saccharomyc
439	37	21.6	564	1	BMRP_CANAL	P28873	candida alb	512	36.5	21.3	575	1	APB3_HUMAN	O96018	homo sapien
440	37	21.6	565	1	COX1_RHOSH	P33517	rhodobacter	513	36.5	21.3	603	1	LEPA_SYNY3	P74751	synechocyst
441	37	21.6	583	1	CRT1_LYCES	P28554	lycopersico	514	36.5	21.3	612	1	OCTC_RAT	P11466	rattus norv
442	37	21.6	622	1	XRC9_HUMAN	Q15287	homo sapien	515	36.5	21.3	618	1	TORA_METH	O27880	methanobact
443	37	21.6	640	1	PPCM_CHICK	P21642	gallus gall	516	36.5	21.3	638	1	SCAD_HUMAN	P51172	homo sapien
444	37	21.6	646	1	Y142_HUMAN	Q14155	homo sapien	517	36.5	21.3	681	1	GFA2_HUMAN	O94808	homo sapien
445	37	21.6	647	1	DPOG_CHICK	Q92076	gallus gall	518	36.5	21.3	681	1	GFA2_MOUSE	Q92229	mus musculu
446	37	21.6	650	1	SCAG_RAT	P37091	rattus norv	519	36.5	21.3	684	1	XYNA_CALSR	O40944	caldicellul
447	37	21.6	728	1	SKI_HUMAN	P12755	homo sapien	520	36.5	21.3	715	1	DCLY_SALTY	Q60002	salmonella
448	37	21.6	733	1	MCMS_MOUSE	P49718	mus musculu	521	36.5	21.3	722	1	Y022_TREPA	O83066	treponema p
449	37	21.6	734	1	MCMS_HUMAN	P33992	homo sapien	522	36.5	21.3	762	1	AOCX_BOVIN	Q29437	bos taurus
450	37	21.6	736	1	KNS1_YEAST	P32350	saccharomyc	523	36.5	21.3	765	1	AOC3_MOUSE	O70423	mus musculu
451	37	21.6	755	1	BGAL_RHIME	O59750	rhizobium m	524	36.5	21.3	803	1	R1R1_CRYPV	O61065	cryptospori
452	37	21.6	816	1	ATX1_HUMAN	P54253	homo sapien	525	36.5	21.3	821	1	SYL_THEPV	P56690	thermus agu
453	37	21.6	825	1	ILAR_HUMAN	P24394	homo sapien	526	36.5	21.3	830	1	EF2_DICDI	P15112	dictyosteli
454	37	21.6	828	1	V2A_TAV	P29035	tomato aspe	527	36.5	21.3	835	1	ULS2_HVSA	P4346	herpesvirus
455	37	21.6	831	1	VPPI_HUMAN	Q93050	homo sapien	528	36.5	21.3	836	1	DPOL_HPBUD	P13162	duck hepati
456	37	21.6	840	1	SYL_BORBU	O51267	borrelia bu	529	36.5	21.3	887	1	HMDH_CRIGR	P03647	cricetulu
457	37	21.6	861	1	SYL_MYCPN	P75258	mycoplasma	530	36.5	21.3	887	1	HMDH_MESAU	P09610	mesocricetu
458	37	21.6	865	1	COX1_ECOLI	P36683	escherichia	531	36.5	21.3	887	1	HMDH_RAT	P51639	rattus norv
459	37	21.6	873	1	CBL_MACEA	Q37370	acanthameob	532	36.5	21.3	888	1	HMDH_HUMAN	O4035	homo sapien
460	37	21.6	896	1	UVR4_RICPR	P22682	mus musculu	533	36.5	21.3	926	1	CHS2_SCHPO	O74756	schizosacch
461	37	21.6	953	1	IDE_DROME	O92cc3	rickettisia	534	36.5	21.3	997	1	YNM3_YEAST	P3920	saccharomyc
462	37	21.6	989	1	IDE_DROME	P22817	drosophila	535	36.5	21.3	1145	1	PR22_YEAST	P24384	saccharomyc
463	37	21.6	1034	1	GCSB_FLAPR	P49362	flaveria pr	536	36.5	21.3	1330	1	XDH_RAT	P22985	rattus norv
464	37	21.6	1034	1	GCSP_FLANR	O49850	flaveria an	537	36.5	21.3	1335	1	XDH_MOUSE	Q00519	mus musculu
465	37	21.6	1034	1	GCSP_FLATR	O49852	flaveria tr	538	36.5	21.3	1358	1	XDH_CHICK	P77990	gallus gall
466	37	21.6	1035	1	GCSP_SOLTU	O49954	solanum tub	539	36.5	21.3	1841	1	RPB1_ARATH	P18616	arabidopsis
467	37	21.6	1037	1	GCSA_FLAPR	P49361	flaveria pr	540	36.5	21.3	1860	1	RPB1_ARATH	P31635	arabidopsis
468	37	21.6	1050	1	ULK1_HUMAN	O75385	homo sapien	541	36.5	21.3	2499	1	MPRI_BOVIN	P08169	bos taurus
469	37	21.6	1073	1	ITA6_HUMAN	P23229	homo sapien	542	36	21.1	75	1	YVBF_VACCC	P20546	vaccinia vi
470	37	21.6	1093	1	SYV_NEUCR	P28350	neurospora	543	36	21.1	81	1	IAC_MACEA	P34953	maccaca fasc
471	37	21.6	1129	1	YB95_YEAST	P38144	saccharomyc	544	36	21.1	105	1	LAC_RABIT	P01847	oryctolagus

545	111	1	SS18_STRVG	P80388 streptomyce	618	36	21.1	413	1	CRF2_XENLA	O42603 xenopus lae
546	120	1	GLHA_EOUAS	Q28365 equus asinu	619	36	21.1	413	1	SVI_CIOIN	Q94425 ciona intes
547	120	1	YB8K_YEAST	P38155 saccharomyc	620	36	21.1	424	1	THIK_HUMAN	P09110 homo sapien
548	123	1	ANG2_BOVIN	P80929 bos taurus	621	36	21.1	431	1	CRF2_MOUSE	Q60748 mus musculu
549	125	1	Y364_AQAE	O66689 aquifex aeo	622	36	21.1	434	1	N075_LUPLU	Q06841 lupinus lut
550	128	1	Y384_HSYMG	Q05102 marek's dis	623	36	21.1	435	1	PTN1_HUMAN	P18031 homo sapien
551	139	1	ANG1_CHICK	P27043 gallus gall	624	36	21.1	441	1	RBL_SYMAL	Q05993 symphoricar
552	140	1	YD23_YEAST	P70749 saccharomyc	625	36	21.1	441	1	RGSE_HUMAN	Q43566 homo sapien
553	141	1	HBAL_PLEWA	P06639 pleurodeles	626	36	21.1	454	1	YXJC_BACSU	P42314 bacillus su
554	145	1	IA02_WHEAT	P16851 triticum ae	627	36	21.1	458	1	RRXB_RAT	P49743 rattus norv
555	150	1	RNK6_GORGO	O46532 gorilla gor	628	36	21.1	459	1	PLSB_ARATH	Q43307 arabidopsis
556	150	1	RNK6_HUMAN	Q93091 homo sapien	629	36	21.1	460	1	HEMO_RABIT	P20058 oryctolagus
557	150	1	RNK6_PANTR	O46525 pan troglod	630	36	21.1	462	1	EF11_CRIGR	P20001 cricetus
558	150	1	RNK6_PONPY	O46526 pongo pygma	631	36	21.1	462	1	EF11_HUMAN	P04720 homo sapien
559	164	1	RB13_RAT	P35286 rattus norv	632	36	21.1	462	1	EF11_MOUSE	P10126 mus musculu
560	201	1	VP24_BDV	P26668 borna disea	633	36	21.1	463	1	EF12_HUMAN	Q05639 homo sapien
561	206	1	YACE_ECOLI	P36679 escherichia	634	36	21.1	463	1	EF12_MOUSE	P27706 mus musculu
562	224	1	NUOB_ERWCA	O85274 erwania car	635	36	21.1	475	1	NCAP_SYNV	P10550 sonchus yel
563	225	1	CRP_MESAU	P49262 mesocricetu	636	36	21.1	476	1	PURA_WHEAT	O24396 triticum ae
564	234	1	YT2_YEAST	P39543 saccharomyc	637	36	21.1	479	1	RBL_ARATH	Q03042 arabidopsis
565	235	1	FRHG_METH	P39498 methanobact	638	36	21.1	479	1	RBL_BRAOL	P48886 brassica ol
566	240	1	VG45_HSV11	Q00111 ictaluriid h	639	36	21.1	487	1	ATF2_HUMAN	P15336 homo sapien
567	244	1	YR01_CABEL	Q10014 caenorhabdi	640	36	21.1	487	1	ATF2_MOUSE	P16951 mus musculu
568	247	1	TRGA_BACSU	P70973 bacillus su	641	36	21.1	487	1	ATF2_RAT	Q00969 rattus norv
569	256	1	ATP6_HANWI	P48879 hansenua w	642	36	21.1	488	1	U2AF_CAEBR	P90727 caenorhabdi
570	265	1	HEM4_HUMAN	P10746 homo sapien	643	36	21.1	490	1	BCHZ_RHOCA	P26179 rhodobacter
571	271	1	HM8_XENLA	P14837 xenopus lae	644	36	21.1	493	1	RBL2_CYACA	P23755 cyanidium c
572	271	1	OX40_RAT	P15725 rattus norv	645	36	21.1	496	1	U2AF_CAEL	P90378 caenorhabdi
573	277	1	HXDB_NOTVI	P31263 notophthalm	646	36	21.1	500	1	TRPE_HELPY	O25869 helicobacte
574	280	1	HXDB_CHICK	P24342 gallus gall	647	36	21.1	501	1	VLI_PAPVE	P11326 europcean el
575	283	1	VP40_MABYM	P35260 marburg vir	648	36	21.1	505	1	SYG_THETH	P56206 thermus aqu
576	289	1	YOLH_BACSU	P42418 bacillus su	649	36	21.1	509	1	HYAL_HUMAN	P38567 homo sapien
577	303	1	ATNB_CANFA	P06583 canis famil	650	36	21.1	509	1	STK_HYDAT	P17713 hydra atten
578	303	1	ATNB_HUMAN	P05026 homo sapien	651	36	21.1	509	1	YD30_MYCTU	Q10641 mycobacteri
579	303	1	ATNB_PIG	P05027 sus scrofa	652	36	21.1	513	1	VLI_PAPVD	P03104 deer papill
580	303	1	ATNB_SHEEP	P05028 ovis aries	653	36	21.1	515	1	COAT_TRSV	Q8894 tobacco rin
581	303	1	PYRD_METH	O27281 methanobact	654	36	21.1	518	1	VLI2_HPV47	P22425 human papil
582	320	1	YDB3_YEAST	P48569 saccharomyc	655	36	21.1	520	1	AMT_BACME	P20845 bacillus me
583	321	1	E13F_HORVU	Q02439 hordeum vul	656	36	21.1	520	1	POLG_HCVH4	Q01404 hepatitis c
584	325	1	RCEM_CHRVI	P51763 chromatium	657	36	21.1	520	1	RXRB_MOUSE	P28704 mus musculu
585	326	1	HEM2_SYNP7	P43087 synechococc	658	36	21.1	533	1	RXRB_HUMAN	P28702 homo sapien
586	326	1	MTM1_CHICK	P08940 gallus gall	659	36	21.1	539	1	FXN_AGRTU	P98055 agrobacteri
587	339	1	PAX9_HUMAN	P55771 homo sapien	660	36	21.1	541	1	YELL_DROME	P09957 drosophila
588	342	1	PAX9_MOUSE	P47242 mus musculu	661	36	21.1	544	1	MYTK_MAIZE	P48190 zea mays (m
589	344	1	MBHS_AZOC	P18190 azotobacter	662	36	21.1	551	1	ATYB_THETU	P19384 thermoanaer
590	344	1	SNPR_STRLI	P43161 streptomyce	663	36	21.1	552	1	PUR6_SCHPO	P15567 schizosacch
591	352	1	THBC_MOUSE	P55104 mus musculu	664	36	21.1	555	1	ODP2_RAT	P08461 rattus norv
592	354	1	CARX_BACST	P54324 bacillus st	665	36	21.1	562	1	GIC4_SOYBN	P02858 glycine max
593	356	1	FL3H_CALCH	Q05963 callistephu	666	36	21.1	585	1	DCEH_YEAST	Q04792 saccharomyc
594	356	1	SMR1_PODAN	Q08142 podospora a	667	36	21.1	587	1	FOLC_MOUSE	P48760 mus musculu
595	358	1	MBHS_RHOCA	P15283 rhodobacter	668	36	21.1	588	1	ADEC_ECOLI	P31441 escherichia
596	360	1	MBHS_ALCEU	P1892 alcaligenes	669	36	21.1	621	1	RPOC_NOSCO	P14563 nostoc comm
597	360	1	MBHS_RHLIV	P18637 rhizobium l	670	36	21.1	632	1	SCAA_XENLA	P51167 xenopus lae
598	360	1	MBHS_RHOGE	P17633 rhodocyclu	671	36	21.1	646	1	PPCK_CHLLI	Q08262 chlorobium
599	363	1	MBHS_ALCHY	P33375 alcaligenes	672	36	21.1	649	1	METF_YEAST	Q04533 saccharomyc
600	363	1	MBHS_BRAJA	P12635 bradyrhizob	673	36	21.1	666	1	BNCH_HELMO	Q92866 heliobacill
601	364	1	FL3H_MALSP	Q06942 malus sp. (674	36	21.1	666	1	ENV_MLVHO	P21436 homulv muri
602	369	1	Y032_SULSO	P95879 sulfolobus	675	36	21.1	669	1	SCAA_HUMAN	P37088 homo sapien
603	370	1	EL3B_PEA	Q03467 pisum sativ	676	36	21.1	675	1	KSC5_ECOLI	P42217 escherichia
604	372	1	MBHS_ECOLI	P19928 escherichia	677	36	21.1	715	1	AT12_HSVIF	P08314 herpes simp
605	374	1	CAN3_PIG	P43368 sus scrofa	678	36	21.1	754	1	CA54_CANFA	Q28247 canis famil
606	379	1	YHM3_YEAST	P38787 saccharomyc	679	36	21.1	756	1	PRLL_HUMAN	Q28274 homo sapien
607	381	1	RODA_HELPY	P56098 helicobacte	680	36	21.1	761	1	AVP3_HORVU	Q06572 hordeum vul
608	388	1	RL3A_ARATH	P17094 arabidopsi	681	36	21.1	769	1	ITB2_BOVIN	P32592 bos taurus
609	388	1	RL3_ORYSA	P35684 oryza sativ	682	36	21.1	769	1	UL06_VZVD	P09302 varicella-z
610	389	1	GSPL_AERHY	P45789 aeromonas h	683	36	21.1	790	1	RIRL_HSVEB	P28846 equine herp
611	389	1	RL3B_ARATH	P22738 arabidopsi	684	36	21.1	795	1	Y348_METJA	Q57794 methanococc
612	399	1	CXAB_CHICK	P36381 gallus gall	685	36	21.1	811	1	OTU_DROME	P10383 drosophila
613	404	1	EAD_EBV	P03191 Epstein-bar	686	36	21.1	833	1	MBP1_YEAST	P39578 saccharomyc
614	411	1	CRF2_HUMAN	Q13324 homo sapien	687	36	21.1	844	1	P8PB_ECOLI	P02619 escherichia
615	411	1	CRF2_RAT	P47866 rattus norv	688	36	21.1	878	1	YD97_SCHPO	Q10337 schizosacch
616	411	1	LAG1_YEAST	P38703 saccharomyc	689	36	21.1	886	1	YFIQ_ECOLI	P76394 escherichia
617	413	1	AGP_ECOLI	P19926 escherichia	690	36	21.1	903	1	DPOL_BPR69	Q3087 bacterioph

691	36	21.1	946	1	YBT6_YEAST	P38250	saccharomyc	764	35.5	20.8	576	1	SBP_CAEL	Q21950	caenorhabdi
692	36	21.1	950	1	UVRA_NEIGO	Q50968	neisseria g	765	35.5	20.8	602	1	LEPA_HELPY	Q25122	helicobacte
693	36	21.1	1023	1	TSCC_PSEAM	P50199	pseudopleur	766	35.5	20.8	604	1	LEPA_HELPY	Q92903	helicobacte
694	36	21.1	1084	1	CRML_YEAST	P30822	saccharomyc	767	35.5	20.8	669	1	FPSL_YEAST	P23900	saccharomyc
695	36	21.1	1114	1	RET_HUMAN	P07949	homo sapien	768	35.5	20.8	772	1	CPTM_HUMAN	Q92523	homo sapien
696	36	21.1	1120	1	DPOL_RCMVM	Q85428	rat cytomeg	769	35.5	20.8	916	1	HMDH_DROME	P14773	drosophila
697	36	21.1	1135	1	VGLM_TSWV1	P36291	tomato spot	770	35.5	20.8	919	1	ENAN_BPK1F	Q04830	bacterioph
698	36	21.1	1153	1	NOS2_HUMAN	P35228	homo sapien	771	35.5	20.8	959	1	MSH1_YEAST	P25846	saccharomyc
699	36	21.1	1170	1	DP3A_TREPA	O83675	treponema p	772	35.5	20.8	1037	1	N120_YEAST	P35729	saccharomyc
700	36	21.1	1284	1	NRCA_CHICK	P30331	gallus gall	773	35.5	20.8	1098	1	PGDR_MOUSE	P05622	mus musc
701	36	21.1	1367	1	IGIR_HUMAN	P08069	homo sapien	774	35.5	20.8	1186	1	SYLC_CAEL	Q09996	caenorhabdi
702	36	21.1	1520	1	ABL_DROME	P00522	drosophila	775	35.5	20.8	1242	1	DPOL_HCMVA	P08546	human cyt
703	36	21.1	1581	1	VGLP_BEV	P23052	berne virus	776	35.5	20.8	1245	1	NIDO_MOUSE	P10493	mus muscu
704	36	21.1	1584	1	UI04_CAEL	P23678	caenorhabdi	777	35.5	20.8	1315	1	CAJH_MOUSE	P39061	mus muscu
705	36	21.1	1669	1	CA14_HUMAN	P04462	homo sapien	778	35.5	20.8	1336	1	SIN3_YEAST	P22579	saccharomyc
706	36	21.1	1669	1	CA14_MOUSE	P04463	mus musc	779	35.5	20.8	2325	1	PGG2_RAT	Q00657	rattus norv
707	36	21.1	1685	1	CA54_HUMAN	P29400	homo sapien	780	35	20.5	37	1	RS15_HELLU	P52820	helix lucor
708	36	21.1	1807	1	VTA2_XENLA	P18709	xenopus lae	781	35	20.5	50	1	SY05_PIG	Q29288	sus scrofa
709	36	21.1	2262	1	RRPL_P12HT	P26676	human parai	782	35	20.5	74	1	UL11_HSVEB	P28982	equine herp
710	36	21.1	2469	1	TEGU_HSVSA	O01056	herpesvirus	783	35	20.5	82	1	VG31_HSVI1	Q00162	ictaluriid h
711	36	21.1	2813	1	VNF_HUMAN	P04275	homo sapien	784	35	20.5	95	1	RS18_RICPR	Q92653	rickettsia
712	36	21.1	3759	1	TRX_DROME	P20659	drosophila	785	35	20.5	113	1	YFRB_PROVU	P20926	proteus vul
713	35.5	20.8	111	1	YK1L_YEAST	P30674	saccharomyc	786	35	20.5	119	1	PA23_NOTSC	P00609	notechis sc
714	35.5	20.8	134	1	N88M_HUMAN	P17568	homo sapien	787	35	20.5	125	1	YGA7_YEAST	P53194	saccharomyc
715	35.5	20.8	142	1	YN8P_YEAST	P53737	saccharomyc	788	35	20.5	127	1	PA2H_XENLA	P41485	xenopus lae
716	35.5	20.8	158	1	SPF4_BOVIN	P81019	bos taurus	789	35	20.5	133	1	TOX5_BORPE	P04981	bordetella
717	35.5	20.8	171	1	IAAD_HORVU	P16643	hordeum vul	790	35	20.5	138	1	PA2C_HUMAN	P39877	homo sapien
718	35.5	20.8	191	1	Y645_METJA	Q58061	methanococc	791	35	20.5	144	1	RS15_HUMAN	P11174	homo sapien
719	35.5	20.8	238	1	ATP6_MYTED	Q00224	mytilus edu	792	35	20.5	145	1	RS15_XIPMA	P70066	xiphophorus
720	35.5	20.8	243	1	GCH1_YEAST	P51601	saccharomyc	793	35	20.5	150	1	CKS1_YEAST	P20486	saccharomyc
721	35.5	20.8	251	1	YAT8_SCHPO	Q10153	schizosacch	794	35	20.5	153	1	IP2K_SOLTU	P01080	solanum tub
722	35.5	20.8	282	1	HUPH_RHILV	P28150	rhizobium l	795	35	20.5	157	1	HSEB_HUMAN	Q14558	homo sapien
723	35.5	20.8	286	1	CH36_DROME	P07182	drosophila	796	35	20.5	164	1	YPHA_ECOLI	P77751	escherichia
724	35.5	20.8	291	1	C61A_HUMAN	P46736	homo sapien	797	35	20.5	165	1	OGT_MYCTU	Q10627	mycobacteri
725	35.5	20.8	292	1	DAPA_ECOLI	P05640	escherichia	798	35	20.5	170	1	Y938_HAEIN	P44079	haemophilus
726	35.5	20.8	293	1	CH36_DROVI	P17111	drosophila	799	35	20.5	179	1	NUOB_AQUAE	O67334	aquifex aeo
727	35.5	20.8	303	1	DDL_THEMEA	P46805	thermotoga	800	35	20.5	182	1	SP23_TENMO	Q27022	tenebrio mo
728	35.5	20.8	303	1	Y85_MYCTU	Q10872	mycobacteri	801	35	20.5	201	1	RB35_HUMAN	Q12886	homo sapien
729	35.5	20.8	314	1	MAG2_HUMAN	P43356	homo sapien	802	35	20.5	207	1	GTP_PIG	P80031	sus scrofa
730	35.5	20.8	318	1	A0X1_MANTIN	Q40294	mangifera i	803	35	20.5	208	1	YPT1_MAIZE	P16976	zea mays (m
731	35.5	20.8	323	1	FLK_RAT	P09760	rattus norv	804	35	20.5	209	1	GTP_RAT	P04906	rattus norv
732	35.5	20.8	344	1	ARGC_THETH	P96136	thermus aqu	805	35	20.5	213	1	VNQN_PAVBO	P07295	bovine parv
733	35.5	20.8	346	1	CDK7_HUMAN	P50613	homo sapien	806	35	20.5	219	1	NUJC_SYNY3	P17082	synecocyst
734	35.5	20.8	346	1	THI3_SCHPO	P36597	schizosacch	807	35	20.5	224	1	GUNX_CLOTM	P15329	clostridium
735	35.5	20.8	376	1	DCOP_CEPAC	P14017	cephalospor	808	35	20.5	229	1	TRYP_SQUAC	P00764	squalus aca
736	35.5	20.8	379	1	DCOP_TRIHA	Q12709	trichoderma	809	35	20.5	231	1	CASB_MOUSE	P10598	mus muscu
737	35.5	20.8	384	1	PURT_BACSU	P39771	bacillus su	810	35	20.5	236	1	CYB_ANGRO	P34861	anguilla ro
738	35.5	20.8	387	1	YV6L_CAEL	Q19978	caenorhabdi	811	35	20.5	243	1	TFS2_ASFB7	P27948	african swi
739	35.5	20.8	393	1	SLAP_BACTI	P38826	bacillus th	812	35	20.5	244	1	PSS9_HUMAN	Q92876	homo sapien
740	35.5	20.8	415	1	CSCB_ECOLI	P30000	escherichia	813	35	20.5	246	1	TRYA_RAT	P32821	rattus norv
741	35.5	20.8	416	1	FDFI_MOUSE	P53798	mus musc	814	35	20.5	246	1	TRYP_MOUSE	P07146	mus muscu
742	35.5	20.8	416	1	FDFI_RAT	Q02769	rattus norv	815	35	20.5	247	1	PSPA_CAVPO	P50403	cavia porce
743	35.5	20.8	417	1	FDFI_HUMAN	P37268	homo sapien	816	35	20.5	247	1	TRX2_BOVIN	Q29463	bos taurus
744	35.5	20.8	436	1	MOBA_ANASP	Q44243	anabaena sp	817	35	20.5	248	1	NUKC_SYNY3	P19050	synecocyst
745	35.5	20.8	448	1	GPT_YEAST	P07286	s udp-n-ace	818	35	20.5	248	1	TRY3_CHICK	Q90629	gallus gall
746	35.5	20.8	461	1	PLSB_PHAVU	Q43822	phaseolus v	819	35	20.5	255	1	HYCC_ECOLI	P16433	escherichia
747	35.5	20.8	473	1	VL2_HPV29	P50800	human papil	820	35	20.5	261	1	HIS6_SYNY3	P74106	synecocyst
748	35.5	20.8	477	1	CB33_YEAST	P35203	saccharomyc	821	35	20.5	262	1	FLGG_CAUCR	Q06172	caulobacter
749	35.5	20.8	477	1	ETV1_HUMAN	P50549	homo sapien	822	35	20.5	262	1	OTCC_NEICI	Q01322	neisseria c
750	35.5	20.8	477	1	ETV1_MOUSE	P41164	mus musc	823	35	20.5	263	1	YHCK_ECOLI	P45427	escherichia
751	35.5	20.8	495	1	THDF_TREPA	O83561	treponema p	824	35	20.5	287	1	RL2_MYCPN	P75577	mycoplasma
752	35.5	20.8	500	1	NIFD_RHICP	P06769	rhizobium s	825	35	20.5	289	1	HOG3_HORVU	P80138	hordeum vul
753	35.5	20.8	503	1	CP3A_MESAU	Q64148	mesocricetu	826	35	20.5	297	1	ATP6_MYCGA	P33251	mycoplasma
754	35.5	20.8	506	1	GAE_HUMAN	P78334	homo sapien	827	35	20.5	299	1	PYRD_ARCFU	Q29513	archaeoglob
755	35.5	20.8	509	1	CX56_CHICK	P29415	gallus gall	828	35	20.5	300	1	CYB_PLAGA	P81376	plasmodium
756	35.5	20.8	518	1	COX1_SQUAC	Q92252	squalus aca	829	35	20.5	300	1	SURL_HUMAN	P81376	plasmodium
757	35.5	20.8	523	1	CBPY_PICPA	P52710	pichia past	830	35	20.5	302	1	YE86_MYCTU	Q15526	homo sapien
758	35.5	20.8	541	1	MYRO_ARATH	P67702	arabidopsis	831	35	20.5	306	1	YCBJ_BACSU	P71766	mycobacteri
759	35.5	20.8	542	1	OPPA_SALTY	P06202	salmonella	832	35	20.5	312	1	PT22_SACKL	P42242	bacillus su
760	35.5	20.8	542	1	XP55_STRLI	P06109	streptomyce	833	35	20.5	313	1	YR24_MYCTU	P13377	saccharomyc
761	35.5	20.8	548	1	MIDL_YEAST	P41921	saccharomyc	834	35	20.5	313	1	YR24_MYCTU	Q10518	mycobacteri
762	35.5	20.8	560	1	RORG_HUMAN	P51421	homo sapien	835	35	20.5	316	1	TAL_HAEIN	Q09356	haemophilus
763	35.5	20.8	572	1	DPY1_HUMAN	Q14194	homo sapien	836	35	20.5	317	1	PLC_LISMO	P34024	listeria mo

837	35	321	1	TA29_TOBAC	P24804	nicotiana t	910	35	20.5	522	1	KTR5_YEAST	P33966	saccharomyc
838	35	321	1	Y4JO_RHISN	P55515	rhizobium s	911	35	20.5	522	1	TSAW_RICTS	P37919	rickettsia
839	35	327	1	HEM2_SYNY3	P77969	synecocyst	912	35	20.5	524	1	LEUL_AQAE	O67862	aquifex aeo
840	35	330	1	E13C_HORVU	Q02126	hordeum vul	913	35	20.5	530	1	UD18_RAT	O84634	rattus norv
841	35	331	1	OTCC_NEIGO	P21302	neisseria g	914	35	20.5	543	1	DAL5_YEAST	P15365	saccharomyc
842	35	333	1	HK32_MOUSE	P97503	mus musculus	915	35	20.5	553	1	Y789_RICPR	O05979	rickettsia
843	35	335	1	RUX_DROME	P50445	drosophila	916	35	20.5	559	1	ECM1_MOUSE	O61508	mus musculus
844	35	345	1	RE12_SCHPO	P40484	schizosacch	917	35	20.5	560	1	LCFA_BAGSU	P64547	bacillus su
845	35	345	1	YLX4_CAEEL	P46500	caenorhabdi	918	35	20.5	566	1	FBLA_HUMAN	P23142	homo sapien
846	35	346	1	H10M_CHICK	Q28256	gallus gall	919	35	20.5	566	1	GUNB_PAELA	P23550	paenibacill
847	35	349	1	LB4D_RABIT	Q28719	oryctolagus	920	35	20.5	569	1	APB3_RAT	O70248	rattus norv
848	35	349	1	TRHB_KLEPN	P06218	klebsiella	921	35	20.5	571	1	APB3_MOUSE	O88888	mus musculus
849	35	352	1	IHBC_HUMAN	P55103	homo sapien	922	35	20.5	571	1	MPPA_NEUCR	P23955	neurospora
850	35	354	1	ARG2_MOUSE	O08691	mus musculus	923	35	20.5	587	1	BARI_YEAST	P12630	saccharomyc
851	35	354	1	MBHS_WOLSU	P31884	wolinella s	924	35	20.5	588	1	RTG2_YEAST	P32608	saccharomyc
852	35	360	1	ARG1_XENLA	Q91553	xenopus lae	925	35	20.5	601	1	FBLB_HUMAN	P23143	homo sapien
853	35	360	1	ARG2_XENLA	Q91554	xenopus lae	926	35	20.5	637	1	SCAA_CHICK	Q92075	gallus gall
854	35	360	1	ARG3_XENLA	Q91555	xenopus lae	927	35	20.5	638	1	KAL_RAT	P14272	rattus norv
855	35	361	1	COOH_RHORU	P31895	rhodospirill	928	35	20.5	639	1	MET7_YEAST	P47164	saccharomyc
856	35	367	1	YL05_YEAST	Q06063	saccharomyc	929	35	20.5	643	1	PCCK_ASCSU	Q05893	ascaris suu
857	35	369	1	GLNA_YEAST	P32288	saccharomyc	930	35	20.5	662	1	COG2_MOUSE	P33434	mus musculus
858	35	372	1	COQ2_YEAST	P32378	saccharomyc	931	35	20.5	662	1	COG2_RAT	P33436	rattus norv
859	35	377	1	TRA7_BACST	Q45620	bacillus st	932	35	20.5	678	1	UL06_HSV2H	P89429	herpes simp
860	35	380	1	CYB_PELSU	O79580	pelomedusa	933	35	20.5	683	1	EFG2_TREPA	O83464	treponema p
861	35	380	1	CYB_XENLA	P00160	xenopus lae	934	35	20.5	683	1	FBLC_HUMAN	P23144	homo sapien
862	35	384	1	BCAT_HUMAN	P54687	homo sapien	935	35	20.5	685	1	FBLC_MOUSE	Q08878	mus musculus
863	35	394	1	GATR_MOUSE	P32336	mus musculus	936	35	20.5	703	1	FBLD_HUMAN	P37888	homo sapien
864	35	396	1	VE2_HPV50	Q80930	human papil	937	35	20.5	705	1	FBLD_MOUSE	Q08879	mus musculus
865	35	398	1	HRCA_CHLPN	Q92850	chlamydia p	938	35	20.5	709	1	GLPK_YEAST	P32190	saccharomyc
866	35	399	1	PRT2_HORVU	P06293	hordeum vul	939	35	20.5	711	1	PT1A_ECOLI	P32670	escherichia
867	35	400	1	ER19_HUMAN	P53602	homo sapien	940	35	20.5	712	1	IRAK_HUMAN	P51617	homo sapien
868	35	404	1	GLGC_MYCTU	O05314	mycobacteri	941	35	20.5	716	1	PERE_MOUSE	P49290	mus musculus
869	35	420	1	MP13_MESAU	P48968	mesocricetu	942	35	20.5	745	1	PERM_HUMAN	P05164	homo sapien
870	35	426	1	GSA_SALTY	P21267	salmonella	943	35	20.5	747	1	FHUA_ECOLI	P06971	escherichia
871	35	428	1	FKH4_MOUSE	Q64733	mus musculus	944	35	20.5	754	1	ECE1_BOVIN	P42891	bos taurus
872	35	432	1	PURA_MYCTU	O08381	mycobacteri	945	35	20.5	754	1	UBP9_YEAST	P39967	saccharomyc
873	35	436	1	RBL_EUGST	P48074	euglena ste	946	35	20.5	821	1	EBN3_EBV	P12977	epstein-bar
874	35	438	1	KMLS_SHEEP	O02827	ovis aries	947	35	20.5	821	1	CAN3_MOUSE	Q04691	mus musculus
875	35	447	1	EF12_DAUCA	P34823	daucus caro	948	35	20.5	821	1	CAN3_RAT	P16259	rattus norv
876	35	448	1	VDR_COTJA	P49701	coturnix co	949	35	20.5	828	1	SYL_RICPR	Q92db1	rickettsia
877	35	451	1	VDR_CHICK	Q42392	gallus gall	950	35	20.5	855	1	DPF6_HUMAN	P42658	homo sapien
878	35	455	1	RBL_LUPNA	P92407	lupinus nan	951	35	20.5	890	1	YOJN_ECOLI	P39838	escherichia
879	35	457	1	RBL_PHECO	P31197	phelline co	952	35	20.5	906	1	CBL_HUMAN	P22681	homo sapien
880	35	458	1	EF1A_ABSGL	P28295	absidia gla	953	35	20.5	907	1	CAPP_MYCLE	P46710	mycobacteri
881	35	459	1	RBL_GEUCH	P28418	geum chiloe	954	35	20.5	914	1	PERT_MOUSE	P35419	mus musculus
882	35	459	1	THDF_BACSU	P25811	bacillus su	955	35	20.5	914	1	PERT_RAT	P14650	rattus norv
883	35	465	1	RBL_COROB	Q32040	cornus obl	956	35	20.5	917	1	VGLB_HSVB2	P12641	bovine herp
884	35	465	1	RBL_DLILN	P28400	dillenia in	957	35	20.5	938	1	NNZ1_HUMAN	Q05586	homo sapien
885	35	465	1	RBL_EPASP	P28412	epacris sp.	958	35	20.5	938	1	NNZ1_MOUSE	P35438	mus musculus
886	35	465	1	RBL_FRAAN	P48703	fragaria an	959	35	20.5	938	1	NNZ1_RAT	P35439	rattus norv
887	35	465	1	RBL_RHOHI	P28447	rhododendro	960	35	20.5	940	1	UVRA_ECOLI	P07671	escherichia
888	35	466	1	RBL_OXADI	P28436	oxalis dill	961	35	20.5	941	1	UVRA_SALTY	P7434	salmonella
889	35	469	1	PPA5_KLULA	P52289	kluyveromyc	962	35	20.5	957	1	UVRA_BAGSU	Q34863	bacillus su
890	35	475	1	RBL_MAGAC	P30732	magnolia ac	963	35	20.5	959	1	DPOL_AAEPE	O03745	aeropyrum p
891	35	475	1	RBL_MAGMA	P30829	magnolia ma	964	35	20.5	964	1	ENV_CAEVG	P31627	caprine art
892	35	481	1	RBL_COFAR	P48594	coffea arab	965	35	20.5	1044	1	ITA8_CHICK	P26009	gallus gall
893	35	482	1	E1BL_ADE12	P04491	human adeno	966	35	20.5	1047	1	MSH3_YEAST	P25336	saccharomyc
894	35	488	1	RBL_ECTSI	P24313	ectocarpus	967	35	20.5	1073	1	ITA6_MOUSE	O61739	mus musculus
895	35	488	1	RBL_PYLIL	P23651	pylaiella l	968	35	20.5	1131	1	YANC_SCHPO	Q10077	schizosacch
896	35	491	1	NUAM_METSE	O47497	metridium s	969	35	20.5	1147	1	KMLS_RABIT	P29294	oryctolagus
897	35	491	1	YOJ6_CAEEL	P34629	caenorhabdi	970	35	20.5	1171	1	NIFJ_KLEPN	P03833	klebsiella
898	35	500	1	DP51_YEAST	P39006	saccharomyc	971	35	20.5	1176	1	KMLS_BOVIN	Q28824	bos taurus
899	35	502	1	GLPK_CAEEL	Q21944	caenorhabdi	972	35	20.5	1213	1	CHS3_CANAL	P30573	candida alb
900	35	503	1	CP3T_PIG	P79401	sus scrofa	973	35	20.5	1220	1	ATCP_HUMAN	P20020	homo sapien
901	35	504	1	VLI_HPV70	P50793	human papil	974	35	20.5	1220	1	ATCP_RABIT	Q00804	oryctolagus
902	35	505	1	VLI_HPV68	P54669	human papil	975	35	20.5	1257	1	RBBI_HUMAN	P29374	homo sapien
903	35	505	1	VLI_HPVME	P27964	human papil	976	35	20.5	1264	1	MOLR_ECOLI	P33345	escherichia
904	35	507	1	CP3S_BOVIN	P79102	bos taurus	977	35	20.5	1266	1	NGCA_CHICK	Q03696	gallus gall
905	35	507	1	VLI_HPV63	O07861	human papil	978	35	20.5	1286	1	SKI2_YEAST	P35207	saccharomyc
906	35	508	1	CATA_HAEIN	P44390	haemophilus	979	35	20.5	1291	1	PIPA_BOVIN	P08487	bos taurus
907	35	510	1	HYAL_WACFA	P38568	macaca fasc	980	35	20.5	1338	1	VGRI_HUMAN	P17948	homo sapien
908	35	519	1	VL2_HPV14	P36749	human papil	981	35	20.5	1390	1	N155_RAT	P37199	rattus norv
909	35	521	1	MET3_YEAST	P08536	saccharomyc	982	35	20.5	1428	1	YA84_SCHPO	Q09773	schizosacch

983 35 20.5 1442 1 CPSA_HUMAN Q10570 homo sapien
984 35 20.5 1442 1 PTC1_CHICK Q30693 gallus gall
985 35 20.5 1444 1 CPSA_BOVIN Q10569 bos taurus
986 35 20.5 1481 1 RPOD_OOSI P49468 odocoileus
987 35 20.5 1529 1 PDRF_YEAST Q04182 saccharomyc
988 35 20.5 1692 1 CYAA_SCHPO P14605 schizosacch
989 35 20.5 1913 1 KMLS_HUMAN Q15746 homo sapien
990 35 20.5 1959 1 AGRI_RAT P25304 rattus norv
991 35 20.5 2228 1 RRPL_SENDZ P06447 sendai viru
992 35 20.5 2324 1 COAC_CHICK P11029 gallus gall
993 35 20.5 2345 1 COAC_RAT P11497 rattus norv
994 35 20.5 2346 1 COAL_HUMAN Q13085 homo sapien
995 35 20.5 2346 1 COAC_SHEEP Q28559 ovis aries
996 35 20.5 3079 1 IRA2_YEAST P19158 saccharomyc
997 35 20.5 3093 1 POLG_BSTV1 Q85730 b genome po
998 34.5 20.2 87 1 PAGD_SALT1 Q56029 salmonella
999 34.5 20.2 110 1 INSI_RAT P01322 rattus norv

ALIGNMENTS

RESULT 1
TSHE_PHOSU STANDARD; PRT; 101 AA.
AC Q62590:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE THYROTROPIN BETA CHAIN (THYROID-STIMULATING HORMONE) (TSH-B)
DE (FRAGMENT).
GN TSHB.
OS Phodopus sungorus (Striped hairy-footed hamster) (Djungarian hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Phodopus.
[1]
RN Rattus norvegicus (Rat).
RP SEQUENCE FROM N.A.
RC TISSUE-PITUITARY;
RX MEDLINE; 96198779.
RA Bockmann J., Bockers T.M., Vennemann B., Niklowitz P., Muller J.,
Witkowski W., Sabel B., Kreutz M.R.;
RT "Short photoperiod-dependent down-regulation of thyrotropin-alpha and
-beta in hamster pars tuberalis-specific cells is prevented by
pinealectomy.";
RT Endocrinology 137:1804-1813(1996).
RL Endocrinology 137:1804-1813(1996).
CC -1- FUNCTION: INDISPENSABLE FOR THE CONTROL OF THYROID STRUCTURE AND
METABOLISM.
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; X90777; CAA62298.1;
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
DR PFAM; PF00007; Cys_knot; 1.
KW Hormone; Glycoprotein.
FT NON_TER 1 1
FT DISULFID 2 88 BY SIMILARITY.
FT DISULFID 10 66 BY SIMILARITY.
FT DISULFID 14 68 BY SIMILARITY.
FT DISULFID 71 78 BY SIMILARITY.
FT CARBOHYD 6 6 POTENTIAL.
SQ SEQUENCE 101 AA; 11353 MW; E5AB4616FC4AE6FF CRC64;

Query Match 93.08; Score 159; DB 1; Length 101;
Best Local Similarity 93.34; Pred. No. 2.5e-16;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 vctyrdfiyrtveipgpcplhvapyfsypva 30
||||| ||||||| ||||||| |||||||
DB 34 VCTYRDFIYRTVEIPGCPHVPVAFSYPVA 63

RESULT 2

TSHE_RAT STANDARD; PRT; 138 AA.
AC P04652:
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE THYROTROPIN BETA CHAIN PRECURSOR (THYROID-STIMULATING HORMONE)
DE (TSH-B).
GN TSHB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
RN Rattus norvegicus (Rat).
RP SEQUENCE FROM N.A.
RX MEDLINE; 86300091.
RA Croyle M.L., Bhattacharya A., Gordon D.F., Maurer R.A.;
RT "Analysis of the organization and nucleotide sequence of the
chromosomal gene for the beta-subunit of rat thyrotropin.";
RL DNA 5:299-304(1986).
[2]
RN Rattus norvegicus (Rat).
RP SEQUENCE FROM N.A.
RC STRAIN-HOLTZMAN;
RX MEDLINE; 84260949.
RA Croyle M.L., Maurer R.A.;
RT "Thyroid hormone decreases thyrotropin subunit mRNA levels in rat
anterior pituitary.";
RL DNA 3:231-236(1984).
[3]
RN Rattus norvegicus (Rat).
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RX MEDLINE; 87109273.
RA Carr F.E., Need L.R., Chin W.W.;
RT "Isolation and characterization of the rat thyrotropin beta-subunit
gene. Differential regulation of two transcriptional start sites by
thyroid hormone.";
RL J. Biol. Chem. 262:981-987(1987).
[4]
RN Rattus norvegicus (Rat).
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RX MEDLINE; 85225532.
RA Chin W.W., Muccini J.A. Jr., Shin L.;
RT "Evidence for a single rat thyrotropin-beta-subunit gene:
thyroidectomy increases its mRNA.";
RL Biochem. Biophys. Res. Commun. 128:1152-1158(1985).
[5]
RN Rattus norvegicus (Rat).
RP SEQUENCE OF 40-138 FROM N.A.
RC STRAIN-WISTAR-IMAMICHI; TISSUE-ANTERIOR PITUITARY;
RA Kato Y., Ezashi T., Hirai T., Kato T.;
RT "Strain difference in nucleotide sequences of rat glycoprotein hormone
subunit cDNAs and gene fragment.";
RL Zool. Sci. 7:877-885(1990).
CC -1- FUNCTION: INDISPENSABLE FOR THE CONTROL OF THYROID STRUCTURE AND
METABOLISM.
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -


```
RL J. Anim. Sci. 74:2104-2111(1996).
RN [3]
RP SEQUENCE OF 21-132.
RX MEDLINE; 76092029.
RA Maguin-Register G., Hennen G., Closset J., Kopeyan C.;
RT "porcine thyrotropin. The amino-acid sequence of the alpha and beta
RT subunits.";
RL Eur. J. Biochem. 61:157-163(1976).
CC -!- FUNCTION: INDISPENSABLE FOR THE CONTROL OF THYROID STRUCTURE AND
CC METABOLISM.
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U39816; AAA93182.1; -.
CC PIR: A01491; TTPGB.
CC PIR; B30339; B30339.
CC HSP; P01233; 1HRP.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_1; 1.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
CC PFAM; PF00007; Cys_knot; 1.
CC Hormone; Glycoprotein; Signal.
CC SIGNAL 1 20
CC CHAIN 21 132 THYROTROPIN BETA CHAIN.
CC PROPEP 133 138
CC DISULFID 22 72 BY SIMILARITY.
CC DISULFID 36 87 BY SIMILARITY.
CC DISULFID 39 125 BY SIMILARITY.
CC DISULFID 47 103 BY SIMILARITY.
CC DISULFID 51 105 BY SIMILARITY.
CC DISULFID 108 115 BY SIMILARITY.
CC CARBOHYD 43 43
CC CONFLICT 42 44 INT -> VNS (IN REF. 3).
CC CONFLICT 57 57 N -> D (IN REF. 3).
CC CONFLICT 109 109 N -> D (IN REF. 3).
CC CONFLICT 129 129 Q -> E (IN REF. 3).
CC SEQUENCE 138 AA; 15761 MW; C75AE4326DAD0377 CRC64;

Query Match 90.1%; Score 154; DB 1; Length 138;
Best Local Similarity 86.7%; Pred. No. 1.9e-15;
Matches 26; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpvhvapyfypva 30
| | | | | | | | | | | | | | | | | | | | | |
DB 71 VCTYRDFMYKTVEIPGCPDHVTPYFSYPVA 100

RESULT 7
TSHE_HORSE STANDARD; PRT; 138 AA.
AC Q28376;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE THYROTROPIN BETA CHAIN PRECURSOR (THYROID-STIMULATING HORMONE)
DE (TSH-B).
GN TSHE.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Perissodactyla; Equidae; Equus.
RN [1]
RP SEQUENCE FROM N.A.

J. Anim. Sci. 74:2104-2111(1996).
RN [3]
RP SEQUENCE OF 21-132.
RX MEDLINE; 76092029.
RA Maguin-Register G., Hennen G., Closset J., Kopeyan C.;
RT "porcine thyrotropin. The amino-acid sequence of the alpha and beta
RT subunits.";
RL Eur. J. Biochem. 61:157-163(1976).
CC -!- FUNCTION: INDISPENSABLE FOR THE CONTROL OF THYROID STRUCTURE AND
CC METABOLISM.
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U39816; AAA93182.1; -.
CC PIR: A01491; TTPGB.
CC PIR; B30339; B30339.
CC HSP; P01233; 1HRP.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_1; 1.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
CC PFAM; PF00007; Cys_knot; 1.
CC Hormone; Glycoprotein; Signal.
CC SIGNAL 1 20
CC CHAIN 21 132 THYROTROPIN BETA CHAIN.
CC PROPEP 133 138
CC DISULFID 22 72 BY SIMILARITY.
CC DISULFID 36 87 BY SIMILARITY.
CC DISULFID 39 125 BY SIMILARITY.
CC DISULFID 47 103 BY SIMILARITY.
CC DISULFID 51 105 BY SIMILARITY.
CC DISULFID 108 115 BY SIMILARITY.
CC CARBOHYD 43 43
CC CONFLICT 42 44 INT -> VNS (IN REF. 3).
CC CONFLICT 57 57 N -> D (IN REF. 3).
CC CONFLICT 109 109 N -> D (IN REF. 3).
CC CONFLICT 129 129 Q -> E (IN REF. 3).
CC SEQUENCE 138 AA; 15761 MW; C75AE4326DAD0377 CRC64;

Query Match 90.1%; Score 154; DB 1; Length 138;
Best Local Similarity 86.7%; Pred. No. 1.9e-15;
Matches 26; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpvhvapyfypva 30
| | | | | | | | | | | | | | | | | | | | | |
DB 71 VCTYRDFMYKTVEIPGCPDHVTPYFSYPVA 100

RESULT 7
TSHE_HORSE STANDARD; PRT; 138 AA.
AC Q28376;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE THYROTROPIN BETA CHAIN PRECURSOR (THYROID-STIMULATING HORMONE)
DE (TSH-B).
GN TSHE.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Perissodactyla; Equidae; Equus.
RN [1]
RP SEQUENCE FROM N.A.

TISSUE-PITUITARY;
Kania S.A., Olchowy T.W., Frank L.A.;
Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: INDISPENSABLE FOR THE CONTROL OF THYROID STRUCTURE AND
CC METABOLISM.
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U51789; AAA96826.1; -.
CC HSP; P01233; 1HRP.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_1; 1.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
CC PFAM; PF00007; Cys_knot; 1.
CC Hormone; Glycoprotein; Signal.
CC SIGNAL 1 20
CC CHAIN 21 132 THYROTROPIN BETA CHAIN.
CC PROPEP 133 138
CC DISULFID 22 72 BY SIMILARITY.
CC DISULFID 36 87 BY SIMILARITY.
CC DISULFID 39 125 BY SIMILARITY.
CC DISULFID 47 103 BY SIMILARITY.
CC DISULFID 51 105 BY SIMILARITY.
CC DISULFID 108 115 BY SIMILARITY.
CC CARBOHYD 43 43
CC SEQUENCE 138 AA; 15663 MW; B03072B4A1C56BA7 CRC64;

Query Match 89.5%; Score 153; DB 1; Length 138;
Best Local Similarity 86.7%; Pred. No. 2.6e-15;
Matches 26; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpvhvapyfypva 30
| | | | | | | | | | | | | | | | | | | | | |
DB 71 VCTYRDFMYKTVEIPGCPDHVTPYFSYPVA 100

RESULT 8
TSHE_BOVIN STANDARD; PRT; 138 AA.
AC P01223;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE THYROTROPIN BETA CHAIN PRECURSOR (THYROID-STIMULATING HORMONE)
DE (TSH-B).
GN TSHE.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
OC Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 84185607.
RA Maurer R.A., Croyle M.L., Donelson J.E.;
RT "The sequence of a cloned cDNA for the beta subunit of bovine
RT thyrotropin predicts a protein containing both NH2 and COOH-terminal
RT extensions.";
RL J. Biol. Chem. 259:5024-5027(1984).
RN [2]
RP SEQUENCE OF 21-132.
RX MEDLINE; 7111428.
RA Liao T.-H., Pierce J.G.;
```



```

RT "The primary structure of bovine thyrotropin. II. The amino acid
RL sequences of the reduced, S-carboxymethyl alpha and beta chains.";
RL J. Biol. Chem. 246:850-865(1971).
[3]
RN PARTIAL SEQUENCE.
RX MEDLINE; 71111427.
RA Shome B., Liao T.-H., Howard S.M., Pierce J.G.;
RT "The primary structure of bovine thyrotropin. I. Isolation and
RL partial sequences of cyanogen bromide and tryptic peptides.";
RL J. Biol. Chem. 246:833-849(1971).
[4]
RN PARTIAL SEQUENCE.
RX MEDLINE; 96239101.
RA Fairlie W.D., Stanton P.G., Hearn T.W.;
RT "The disulphide bond structure of thyroid-stimulating hormone beta-
RL subunit.";
RL Biochem. J. 314:449-455(1996).
CC -1- FUNCTION: INDISPENSABLE FOR THE CONTROL OF THYROID STRUCTURE AND
CC METABOLISM.
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K01939; AAA30796.1; -.
CC PIR; A01490; TTEOB.
CC HSP; P01233; LHRP.
CC DR
CC DR
CC PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
CC DR
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
CC DR
CC PFAM; PF00007; Cys_knot; 1.
CC Hormone; Glycoprotein; Signal.
CC KW
CC SIGNAL 1 20
CC CHAIN 21 132 THYROTROPIN BETA CHAIN.
CC FT
CC PROPEP 133 138
CC FT
CC FT DISULFID 22 72
CC FT DISULFID 36 87
CC FT DISULFID 39 125
CC FT DISULFID 47 103
CC FT DISULFID 51 105
CC FT DISULFID 108 115
CC FT CARBOHYD 43 43
CC SQ SEQUENCE 138 AA; 15624 MW; 42D783B7C0E2EB98 CRC64;

Query Match 88.3%; Score 151; DB 1; Length 138;
Best Local Similarity 83.3%; Pred. No. 5.1e-15;
Matches 25; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 vctyrdgfvrtveipgcpnlhvapyfsypva 30
DB 71 VCYTRDPGRTAEIPGCPRHVTPIFSIPVA 100
|||||||:|||||||:|||||||
|||||||:|||||||:|||||||

RESULT 9
TSBH_CHICK
ID TSBH_CHICK STANDARD; PRT; 134 AA.
AC O57340;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE THYROTROPIN BETA CHAIN PRECURSOR (THYROID-STIMULATING HORMONE)
DE (TSB-B).
TSBH.
OS Gallus gallus (Chicken).
```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
OC Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=ANTERIOR PITUITARY;
RX MEDLINE; 97392782.
RA Gregory C.C., Porter T.E.;
RT "Cloning and sequence analysis of a cDNA for the beta subunit of
RL chicken thyroid-stimulating hormone.";
RL Gen. Comp. Endocrinol. 107:182-190(1997).
CC -1- FUNCTION: INDISPENSABLE FOR THE CONTROL OF THYROID STRUCTURE AND
CC METABOLISM (BY SIMILARITY).
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF033495; AAB88127.1; -.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
DR PFAM; PF00007; Cys_knot; 1.
KW Hormone; Glycoprotein; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 134 THYROTROPIN BETA CHAIN.
FT DISULFID 22 72 BY SIMILARITY.
FT DISULFID 36 87 BY SIMILARITY.
FT DISULFID 39 125 BY SIMILARITY.
FT DISULFID 47 103 BY SIMILARITY.
FT DISULFID 51 105 BY SIMILARITY.
FT DISULFID 108 115 BY SIMILARITY.
FT CARBOHYD 43 43 POTENTIAL.
FT SEQUENCE 134 AA; 15064 MW; 3D2429DB4EAC07EB CRC64;

Query Match 69.6%; Score 119; DB 1; Length 134;
Best Local Similarity 63.3%; Pred.No. 2e-10;
Matches 19; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 vctyrdfyrtveipgclhvayfsypva 30
||||: |: ||||| ||: |||||
Db 71 VCTYKEMFYQTALIPGCPHPTIPYSYPVA 100

RESULT 10
TSHE_ANGAN
ID TSHB_ANGAN STANDARD; PRT; 147 AA.
AC Q08127;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE THYROTROPIN BETA CHAIN PRECURSOR ("PHYROID-STIMULATING HORMONE")
DE ("TSH-B").
OS Anguilla anguilla (European freshwater eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Anguilliformes; Anguillidae; Anguilla.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 94319976.
RA Salmon C., Marchelidon J., Fontaine Y.A., Huet J.-C., Querat B.;
RT "Cloning and sequence of thyrotropin beta subunit of a teleost fish:
RT the eel (Anguilla anguilla L.).";
RL C. R. Acad. Sci., III, Sci. Vie 316:749-753(1993).
[2]
RN SEQUENCE OF 21-43.

```


CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D14692; BAA03518.1; -
DR PIR: A48194; A48194.
DR HSSP: P01233; IHRP.
DR PRINTS: PR00438; GFCYSKNOT.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
DR PFAM: PF00007; Cys_knot; 1.
KW Hormone; Glycoprotein; Signal.
FT SIGNAL 1 20
FT CHAIN 21 147 THYROTROPIN BETA CHAIN.
FT DISULFID 22 72 BY SIMILARITY.
FT DISULFID 36 87 BY SIMILARITY.
FT DISULFID 39 127 BY SIMILARITY.
FT DISULFID 47 103 BY SIMILARITY.
FT DISULFID 51 105 BY SIMILARITY.
FT DISULFID 108 115 BY SIMILARITY.
FT CARBOHYD 43 43 POTENTIAL.
SQ SEQUENCE 147 AA; 16440 MW; 0F1642BA9FCA35BA CRC64;

Query Match 63.7%; Score 109; DB 1; Length 147;
Best Local Similarity 65.5%; Pred. No. 6.3e-09;
Matches 19; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 ctyrdfyrtveipgcpvhvafsyypva 30
DB 72 CTYDQVEYRTVILPGCPHLANPLFTYTPVA 100

RESULT 13
GTH2_THUOB
ID GTH2_THUOB STANDARD; PRT; 115 AA.
AC P37206;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GONADOTROPIN BETA-II CHAIN (GTH-II).
OS Thunnus obesus (Bigeye tuna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Perciformes; Scombroidei; Scombridae; Thunnus.
RN [1]
RP SEQUENCE.
RC TISSUE-PITUITARY;
RX MEDLINE; 94186287.
RA Okada T., Kawazoe I., Kimura S., Sasamoto Y., Aida K., Kawauchi H.;
RT "Purification and characterization of gonadotropin I and II from
RT pituitary glands of tuna (Thunnus obesus).";
RL Int. J. Pept. Protein Res. 43:69-80(1994).
CC -|- FUNCTION: INVOLVED IN GAMETOGENESIS AND STEROIDOGENESIS.
CC -|- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -|- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
DR HSSP: P01233; IHRP.
DR PRINTS: PR00438; GFCYSKNOT.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
DR PFAM: PF00007; Cys_knot; 1.
KW Hormone; Glycoprotein.
FT DISULFID 6 54 BY SIMILARITY.
FT DISULFID 20 69 BY SIMILARITY.
FT DISULFID 23 107 BY SIMILARITY.
FT DISULFID 31 85 BY SIMILARITY.
FT DISULFID 35 87 BY SIMILARITY.
FT DISULFID 90 97 BY SIMILARITY.

FT CARBOHYD 10 10
SQ SEQUENCE 115 AA; 12992 MW; 8721C07836F5CAE9 CRC64;

Query Match 62.6%; Score 107; DB 1; Length 115;
Best Local Similarity 63.3%; Pred. No. 9.4e-09;
Matches 19; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 vctyrdfyrtveipgcpvhvafsyypva 30
DB 53 VCTYRDFYRTVEIPGCPVGVDPVTYTPVA 82

RESULT 14
GTHB_HYPMO
ID GTHB_HYPMO STANDARD; PRT; 141 AA.
AC P37038;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GONADOTROPIN BETA CHAIN PRECURSOR.
OS Hypophthalmichthys molitrix (Silver carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinidae; Leuciscinae; Hypophthalmichthys.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE-PITUITARY;
RX MEDLINE; 90236229.
RA Chang Y.S., Huang C.J., Huang F.-L., Liu C.S., Lo T.-B.;
RT "Purification, characterization, and molecular cloning of
RT gonadotropin subunits of silver carp (Hypophthalmichthys molitrix).";
RL Gen. Comp. Endocrinol. 78:23-33(1990).
CC -|- FUNCTION: INVOLVED IN GAMETOGENESIS AND STEROIDOGENESIS.
CC -|- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -|- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
DR PIR: B60626; B60626.
DR HSSP: P01233; IHRP.
DR PRINTS: PR00438; GFCYSKNOT.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
DR PFAM: PF00007; Cys_knot; 1.
KW Hormone; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 139 GONADOTROPIN BETA CHAIN.
FT PROPEP 140 141
FT DISULFID 30 78 BY SIMILARITY.
FT DISULFID 44 93 BY SIMILARITY.
FT DISULFID 47 131 BY SIMILARITY.
FT DISULFID 55 109 BY SIMILARITY.
FT DISULFID 59 111 BY SIMILARITY.
FT DISULFID 114 121 BY SIMILARITY.
FT CARBOHYD 34 34 POTENTIAL.
SQ SEQUENCE 141 AA; 15856 MW; A42C48FE983EEA46 CRC64;

Query Match 59.1%; Score 101; DB 1; Length 141;
Best Local Similarity 60.0%; Pred. No. 8.6e-08;
Matches 18; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 vctyrdfyrtveipgcpvhvafsyypva 30
DB 77 VCTYRDFYRTVEIPGCPVGVDPHTYTPVA 106

RESULT 15
GTHB_CYPCA
ID GTHB_CYPCA STANDARD; PRT; 144 AA.
AC P01235;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)

RX MEDLINE: 73090987.
 RA Shone B., Parlow A.F.;
 RT "The primary structure of the hormone-specific, beta subunit of human
 RT pituitary luteinizing hormone (hLH).";
 RN J. Clin. Endocrinol. Metab. 36:618-621(1973).
 RN [4]
 RP PRELIMINARY PARTIAL SEQUENCE.
 RX MEDLINE: 73221227.
 RA Closset J., Hennen G., Lequin R.M.;
 RT "Human luteinizing hormone. The amino acid sequence of the
 RT subunit.";
 RN FEBS Lett. 29:97-100(1973).
 RN [5]
 RP STRUCTURE OF CARBOHYDRATE.
 RX MEDLINE: 91122088.
 RA Weissshaar G., Hiyama J., Renwick A.G.C., Nintz M.;
 RT "NMR investigations of the N-linked oligosaccharides at individual
 RT glycosylation sites of human lutropin.";
 RN Eur. J. Biochem. 195:257-268(1991).
 RN [6]
 RP STRUCTURE BY NMR OF 58-77.
 RX MEDLINE: 92357029.
 RA Keutmann H.T., Hua Q.-X., Weiss M.A.;
 RT "Structure of a receptor-binding fragment from human luteinizing
 RT hormone beta-subunit determined by [1H]- and [15N]nuclear magnetic
 RT resonance spectroscopy.";
 RN Mol. Endocrinol. 6:904-913(1992).
 RN [7]
 RP VARIANT ARG-74.
 RX MEDLINE: 92085985.
 RA Weiss J., Axelrod L., Whitcomb R.W., Harris P.E., Crowley W.F.,
 RA Jameson J.L.;
 RT "Hypogonadism caused by a single amino acid substitution in the beta
 RT subunit of luteinizing hormone.";
 RN New Engl. J. Med. 326:179-183(1992).
 CC -!- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
 CC -!- THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
 CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
 CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
 CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
 CC -!- TISSUE SPECIFICITY: PITUITARY.
 CC -!- DISEASE: DEFECTS IN LHB ARE A CAUSE OF HYPOGONADISM WHICH IS
 CC CHARACTERIZED BY INFERTILITY AND PSEUDOHERMAPHRODITISM.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X00264; CAA25067.1; -.
 DR EMBL: S71273; AAD14960.1; ALT_SEQ.
 DR PIR: A01497; UTHUB.
 DR HSP: P01233; 1HRP.
 DR MIM: 152780; -.
 DR PRINTS: PR00438; GFCYSKNOT.
 DR PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.
 DR PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.
 DR PFAM: PF00007; Cys_knot; 1.
 KW Hormone; Glycoprotein; Signal; Pseudohermaphroditism;
 KW Disease mutation.
 FT SIGNAL 1 20
 FT CHAIN 21 141 LUTROPIN BETA CHAIN.
 FT DISULFID 29 77 BY SIMILARITY.
 FT DISULFID 43 92 BY SIMILARITY.
 FT DISULFID 46 130 BY SIMILARITY.
 FT DISULFID 54 108 BY SIMILARITY.
 FT DISULFID 58 110 BY SIMILARITY.
 FT DISULFID 113 120 BY SIMILARITY.

FT CARBOHYD 50 50 Q -> R (IN HYPOGONADISM; LACK OF
 FT VARIANT 74 74 RECEPTOR-BINDING).
 FT /FTID-VAR_003189.
 FT E -> Q (IN REF. 2).
 FT CONFLICT 39 39 MISSING (IN REF. 2).
 FT CONFLICT 76 76 MISSING (IN REF. 2).
 FT CONFLICT 132 135 HPOL -> POH (IN REF. 2).
 SQ SEQUENCE 141 AA; 15345 MW; E411766253113F7C CRC64;

 Query Match 55.6%; Score 95; DB 1; Length 141;
 Best Local Similarity 53.3%; Pred. No. 6.3e-07;
 Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

 QY 1 vctyrdfiyrtveipgcpvhvapyfsypva 30
 ||||| : : : |||| | | : ||||
 DB 76 VCTYRDVFESIRLPGCGRGVDPVVSFPVA 105

 RESULT 20
 FSHB_STRCA STANDARD; PRT; 106 AA.
 ID FSHB_STRCA
 AC P80663;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE FOLLITROPIN BETA CHAIN (FOLLICLE-STIMULATING HORMONE) (FSH-B).
 OS Struthio camelus (Ostrich).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;
 OC Palaeognathae; Struthioniformes; Struthionidae; Struthio.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE: 97025333.
 RA Koide Y., Pakkoff H., Kawauchi H.;
 RT "Complete amino acid sequences of follitropin and lutropin in the
 RT ostrich, Struthio camelus.";
 RL Eur. J. Biochem. 240:262-267(1996).
 CC -!- FUNCTION: STIMULATES DEVELOPMENT OF FOLLICLE AND SPERMATOGENESIS
 CC IN THE REPRODUCTIVE ORGANS.
 CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
 CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
 CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
 CC FAMILY.
 DR HSP: P01233; 1HRP.
 DR PRINTS: PR00438; GFCYSKNOT.
 DR PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.
 DR PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.
 DR PFAM: PF00007; Cys_knot; 1.
 KW Hormone; Glycoprotein.
 FT DISULFID 1 49 BY SIMILARITY.
 FT DISULFID 15 64 BY SIMILARITY.
 FT DISULFID 18 102 BY SIMILARITY.
 FT DISULFID 26 80 BY SIMILARITY.
 FT DISULFID 30 82 BY SIMILARITY.
 FT DISULFID 85 92 BY SIMILARITY.
 FT CARBOHYD 5 5 POTENTIAL.
 FT CARBOHYD 22 22 POTENTIAL.
 SQ SEQUENCE 106 AA; 11947 MW; 85AD110D1C5CD99B CRC64;

 Query Match 55.0%; Score 94; DB 1; Length 106;
 Best Local Similarity 51.7%; Pred. No. 6.5e-07;
 Matches 15; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

 QY 2 cttyrdfiyrtveipgcpvhvapyfsypva 30
 ||::: : ||::| | : |||||
 DB 49 CTFKEVYVETVKIPGCRDHAESLYSPVA 77

 RESULT 21
 FSHB_MOUSE
 ID FSHB_MOUSE STANDARD; PRT; 130 AA.


```
AC Q06087;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE FOLLITROPIN BETA CHAIN PRECURSOR (FOLLICLE-STIMULATING HORMONE)
DE (FSH-B).
GN FSHB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RA Kumar R.T.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STIMULATES DEVELOPMENT OF FOLLICLE AND SPERMATOGENESIS
CC IN THE REPRODUCTIVE ORGANS.
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U12932; AAA92804.1; -.
CC HSP; P01233; IHRP.
CC MGD; MGI:95582; FSHB.
CC PRINTS; PR00438; GFCYSKNOT.
CC PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
CC PFAM; PF00007; Cys_knot; 1.
CC Hormone; Glycoprotein; Signal.
CC SIGNAL 1 20 BY SIMILARITY.
CC CHAIN 21 130 FOLLITROPIN BETA CHAIN.
CC FT DISULFID 22 70 BY SIMILARITY.
CC FT DISULFID 36 85 BY SIMILARITY.
CC FT DISULFID 39 123 BY SIMILARITY.
CC FT DISULFID 47 101 BY SIMILARITY.
CC FT DISULFID 51 103 BY SIMILARITY.
CC FT DISULFID 106 113 BY SIMILARITY.
CC FT CARBOHYD 26 26 POTENTIAL.
CC FT CARBOHYD 43 43 POTENTIAL.
CC SEQUENCE 130 AA; 14919 MW; 7F9C28C2E34AC161 CRC64;

Query Match 54.4%; Score 93; DB 1; Length 130;
Best Local Similarity 46.7%; Pred. No. 1.1e-06;
Matches 14; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcpplhvayfsypva 30
||||| : ||| : ||| : |||
Db 69 VCTEKELVYETVRLPGCARHSDSLYTPVA 98

RESULT 22
GTH2_ONCMA STANDARD; PRT; 142 AA.
AC P48253;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GONADOTROPIN BETA-II CHAIN PRECURSOR (GTH-II).
DE ONCORTHYRINCHUS MASOU (Cherry salmon).
OS ONCORTHYRINCHUS MASOU (Cherry salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
OC Salmoniformes; Salmonidae; Oncorhynchus.
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC TISSUE-PITUITARY;
RA MEDLINE; 94197893.
RX Kato Y., Gen K., Maruyama O., Tomizawa K., Kato T.;
RA "Molecular cloning of cDNAs encoding two gonadotrophin beta subunits
RT (GTH-I beta and -II beta) from the masu salmon, Oncorhynchus masou:
RT rapid divergence of the GTH-I beta gene.";
RL J. Mol. Endocrinol. 11:275-282(1993).
CC -1- FUNCTION: INVOLVED IN GAMETOGENESIS AND STEROIDOGENESIS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S69276; AAB30424.1; -.
CC HSP; P01233; IHRP.
CC PRINTS; PR00438; GFCYSKNOT.
CC PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
CC PFAM; PF00007; Cys_knot; 1.
CC Hormone; Glycoprotein; Signal.
CC SIGNAL 1 23 BY SIMILARITY.
CC CHAIN 24 142 GONADOTROPIN BETA-II CHAIN.
CC FT DISULFID 29 77 BY SIMILARITY.
CC FT DISULFID 43 92 BY SIMILARITY.
CC FT DISULFID 46 130 BY SIMILARITY.
CC FT DISULFID 54 108 BY SIMILARITY.
CC FT DISULFID 58 110 BY SIMILARITY.
CC FT DISULFID 113 120 BY SIMILARITY.
CC FT CARBOHYD 33 33 POTENTIAL.
CC SEQUENCE 142 AA; 15884 MW; 70A3906EAB3FF CRC64;

Query Match 54.4%; Score 93; DB 1; Length 142;
Best Local Similarity 53.3%; Pred. No. 1.2e-06;
Matches 16; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcpplhvayfsypva 30
||||| : ||| : ||| : |||
Db 76 VCTYRDVRYEMIRLPDCPPWDPHVTPVA 105

RESULT 23
FSHB_HORSE STANDARD; PRT; 118 AA.
AC P01226;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE FOLLITROPIN BETA CHAIN (FOLLICLE-STIMULATING HORMONE) (FSH-B).
GN FSHB.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Perissodactyla; Equidae; Equus.
RN [1]
RP SEQUENCE.
RX MEDLINE; 78218213.
RA Fujiki Y., Rathnam P., Saxena B.B.;
RT "Amino acid sequence of the beta-subunit of the follicle-stimulating
RT hormone from equine pituitary glands.";
RL J. Biol. Chem. 253:5363-5368(1978).
CC -1- FUNCTION: STIMULATES DEVELOPMENT OF FOLLICLE AND SPERMATOGENESIS
CC IN THE REPRODUCTIVE ORGANS.
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC
```


[illegible]

RA Closset J., Maghuin-Rogister G., Hennen G., Strosber

RA Closser


```
RT "porcine follitropin. The amino-acid sequence of the beta subunit.";
RL Eur. J. Biochem. 86:115-120(1978).
CC -!- FUNCTION: STIMULATES DEVELOPMENT OF FOLLICLE AND SPERMATOGENESIS
CC IN THE REPRODUCTIVE ORGANS.
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D00621; BAA00499.1; -.
DR EMBL; M35676; AAA31039.1; -.
DR PIR; A01496; FTGGB.
DR PIR; A48169; A48169.
DR HSSP; P01233; IHRP.
DR PRINTS; PR00438; GFCYSKNOT.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
DR PFAM; PF00007; Cys_knot; 1.
KW Hormone; Glycoprotein; Signal.
FT SIGNAL 1 20
FT CHAIN 21 129 FOLLITROPIN BETA CHAIN.
FT DISULFID 21 69 BY SIMILARITY.
FT DISULFID 35 84 BY SIMILARITY.
FT DISULFID 38 122 BY SIMILARITY.
FT DISULFID 46 100 BY SIMILARITY.
FT DISULFID 50 102 BY SIMILARITY.
FT DISULFID 105 112 BY SIMILARITY.
FT CARBOHYD 25 25 PROBABLE.
FT CARBOHYD 42 42 PROBABLE.
FT CONFLICT 32 36 KECN -> VKLT (IN REF. 3).
FT CONFLICT 51 52 YT -> TTG (IN REF. 3).
FT CONFLICT 71 72 FK -> YR (IN REF. 3).
FT CONFLICT 125 125 S -> G (IN REF. 3).
SQ SEQUENCE 129 AA; 14605 MW; 20BBCBEDF209E1EA CRC64;

Query Match 53.8%; Score 92; DB 1; Length 129;
Best Local Similarity 44.8%; Pred. No. 1.6e-06;
Matches 13; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 2 ctvrdfiyrtveipgcpvhvafsyypva 30
Db 69 CTFKELVETVKVPGCAHADSLSLYTPVA 97
||||: |||:|||| |:::|||||

RESULT 26
FSHB_SHEEP STANDARD; PRT; 129 AA.
AC P01227;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE FOLLITROPIN BETA CHAIN PRECURSOR (FOLLICLE-STIMULATING HORMONE)
DE (FSH-B).
GN FSHB.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92029621.
RA Guzman K., Miller C.D., Phillips C.L., Miller W.L.;
RT "The gene encoding ovine follicle-stimulating hormone beta:
```

```
RT isolation, characterization, and comparison to a related ovine
RT genomic sequence.";
RL DNA Cell Biol. 10:593-601(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89366671.
RA Mountford P.S., Bello P.A., Brandon M.R., Adams T.E.;
RT "Cloning and DNA sequence analysis of the cDNA for the precursor of
RT ovine follicle stimulating hormone beta-subunit.";
RL Nucleic Acids Res. 17:6391-6391(1989).
RN [3]
RP SEQUENCE OF 20-129.
RX MEDLINE; 82113053.
RA Sairam M.R., Seidah N.G., Chretien M.;
RT "Primary structure of the ovine pituitary follitropin beta-subunit.";
RL Biochem. J. 197:541-552(1981).
CC -!- FUNCTION: STIMULATES DEVELOPMENT OF FOLLICLE AND SPERMATOGENESIS
CC IN THE REPRODUCTIVE ORGANS.
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S64745; AAB20317.1; -.
DR EMBL; X15493; CAA33516.1; -.
DR PIR; A01495; FTSHB.
DR PIR; A40410; A40410.
DR PIR; S05316; S05316.
DR HSSP; P01233; IHRP.
DR PRINTS; PR00438; GFCYSKNOT.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
DR PFAM; PF00007; Cys_knot; 1.
KW Hormone; Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 129 FOLLITROPIN BETA CHAIN.
FT DISULFID 21 69 BY SIMILARITY.
FT DISULFID 35 84 BY SIMILARITY.
FT DISULFID 38 122 BY SIMILARITY.
FT DISULFID 46 100 BY SIMILARITY.
FT DISULFID 50 102 BY SIMILARITY.
FT DISULFID 105 112 BY SIMILARITY.
FT CARBOHYD 25 25 PROBABLE.
FT CARBOHYD 42 42 PROBABLE.
FT CONFLICT 68 68 A -> T (IN REF. 3).
FT CONFLICT 107 107 R -> T (IN REF. 3).
FT CONFLICT 128 129 RE -> ERZ (IN REF. 3).
SQ SEQUENCE 129 AA; 14669 MW; 83D76DCDC971EF40 CRC64;

Query Match 53.8%; Score 92; DB 1; Length 129;
Best Local Similarity 44.8%; Pred. No. 1.6e-06;
Matches 13; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 2 ctvrdfiyrtveipgcpvhvafsyypva 30
Db 69 CTFKELVETVKVPGCAHADSLSLYTPVA 97
||||: |||:|||| |:::|||||

RESULT 27
FSHB_TRIVU STANDARD; PRT; 129 AA.
ID FSHB_TRIVU
AC O46430;
DT 15-DEC-1998 (Rel. 37, Created)
```



```

RESULT 29
LSHB_MOUSE STANDARD; PRT; 105 AA.
AC 009108;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE) (LSH-B) (LH-B)
DE (FRAGMENT).
GN LHB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 x CBA; TISSUE=ANTERIOR PITUITARY;
RA Brown P., Brooks J., McNeilly J.R., McNeilly A.S.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
CC THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y10418; CAA71445.1; -
DR HSSP; P01233; LHBP.
DR MGD; MGI:96782; LHB.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; FALSE_NEG.
DR PFAM; PF00007; Cys_knot; 1.
KW Hormone; Signal; Glycoprotein.
FT NON_TER 1
FT SIGNAL <1 3 BY SIMILARITY.
FT CHAIN 4 >105 LUTROPIN BETA CHAIN.
FT DISULFID 12 60 BY SIMILARITY.
FT DISULFID 26 75 BY SIMILARITY.
FT DISULFID 37 91 BY SIMILARITY.
FT DISULFID 41 93 BY SIMILARITY.
FT DISULFID 96 103 BY SIMILARITY.
FT CARBOHYD 16 16 POTENTIAL.
FT NON_TER 105 105
SQ SEQUENCE 105 AA; 11095 MW; 5DB7831704C6DEA6 CRC64;

Query Match 53.2%; Score 91; DB 1; Length 105;
Best Local Similarity 53.3%; Pred. No. 1.7e-06;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgclhvpafyspva 30
||||| : : | : ||| | : | : |||
Db 59 VCTYRELRFASVRLPGCPGVDPIVSFPA 88

RESULT 30
FSHB_HUMAN STANDARD; PRT; 129 AA.
AC P01225;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE FOLLITROPIN BETA CHAIN PRECURSOR (FOLLICLE-STIMULATING HORMONE)

```

DE (FSH-B).

GN FSHB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Homiidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 87246070.

RA Watkins P.C., Eddy R., Beck A.K., Vellucci V., Leverone B.,

RA Tanzi R.E., Gusella J.F., Shows T.B.;

RT "DNA sequence and regional assignment of the human

RT follicle-stimulating hormone beta-subunit gene to the short arm of

RT human chromosome 11.";

RL DNA 6:205-212(1987).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE; 89174746.

RA Keene J.L., Matzuk M.M., Otani T., Fauser B.C.J.M., Galway A.B.,

RA Hsueh A.J.W., Boime I.;

RT "Expression of biologically active human follitropin in Chinese

RT hamster ovary cells.";

RL J. Biol. Chem. 264:4769-4775(1989).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE; 89014411.

RA Jameson J.L., Becker C.B., Lindell C.M., Habener J.F.;

RT "Human follicle-stimulating hormone beta-subunit gene encodes

RT multiple messenger ribonucleic acids.";

RL Mol. Endocrinol. 2:806-815(1988).

RN [4]

RP SEQUENCE OF 19-129.

RX MEDLINE; 89351581.

RA Shome B., Parlow A.F., Liu W.K., Nahm H.S., Wen T., Ward D.N.;

RT "A reevaluation of the amino acid sequence of human follitropin beta-

RT subunit.";

RL J. Protein Chem. 7:325-339(1988).

RN [5]

RP PRELIMINARY SEQUENCE OF 19-129.

RX MEDLINE; 76120602.

RA Saxena B.B., Rathnam P.;

RT "Amino acid sequence of the beta subunit of follicle-stimulating

RT hormone from human pituitary glands.";

RL J. Biol. Chem. 251:993-1005(1976).

RN [6]

RP PRELIMINARY SEQUENCE OF 19-129.

RX MEDLINE; 74262938.

RA Shome B., Parlow A.F.;

RT "Human follicle stimulating hormone: first proposal for the amino

RT acid sequence of the hormone-specific, beta subunit (hFSHb).";

RL J. Clin. Endocrinol. Metab. 39:203-205(1974).

RN [7]

RP PRELIMINARY ASSIGNMENT OF DISULFIDE BONDS.

RX MEDLINE; 81021713.

RA Fujiki Y., Rathnam P., Saxena B.B.;

RT "Studies on the disulfide bonds in human pituitary

RT follicle-stimulating hormone.";

RL Biochim. Biophys. Acta 624:428-435(1980).

CC -1- FUNCTION: STIMULATES DEVELOPMENT OF FOLLICLE AND SPERMATOGENESIS

CC IN THE REPRODUCTIVE ORGANS.

CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA

CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,

CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.

CC -1- PHARMACEUTICAL: AVAILABLE UNDER THE NAMES GONAL-F OR METRODIN HP

CC (SERONO) AND PUREGON (ORGANON).

CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN

CC FAMILY.

CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)


```
CC THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y16265; CA076146.1; -.
CC PRINTS; PR00438; GFCYSKNOT.
CC PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
CC PFAM; PF00007; Cys_knot; 1.
CC Hormone; Glycoprotein; Signal.
CC SIGNAL 1 20 BY SIMILARITY.
CC CHAIN 21 169 LUTROPIN/CHORIOGONADOTROPIN BETA CHAIN.
CC DISULFID 29 77 BY SIMILARITY.
CC DISULFID 43 92 BY SIMILARITY.
CC DISULFID 46 130 BY SIMILARITY.
CC DISULFID 54 108 BY SIMILARITY.
CC DISULFID 58 110 BY SIMILARITY.
CC DISULFID 113 120 BY SIMILARITY.
CC CARBOHYD 33 33 POTENTIAL.
CC CARBOHYD 158 158 POTENTIAL.
CC SEQUENCE 169 AA; 17824 MW; 322DF724EAA93E9 CRC64;

Query Match 52.6%; Score 90; DB 1; Length 169;
Best Local Similarity 50.0%; Pred. No. 4e-06;
Matches 15; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcpbhvafsyypva 30
||||| : : : ||||| | | |||||
Db 76 VCTYRELREASIRLPGCPGVDPWVSFPVA 105

RESULT 35
LSHB_HORSE
ID LSHB_HORSE STANDARD; PRT; 169 AA.
AC P08751; P01234;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE LUTROPIN/CHORIOGONADOTROPIN BETA CHAIN PRECURSOR (LSH-B/CG-B)
DE (LUTEINIZING HORMONE BETA SUBUNIT).
LHB.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Perissodactyla; Equidae; Equus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92357035.
RA Sherman G.B., Wolfe M.W., Farmerie T.A., Clay C.M.,
RA Threadgill D.S., Sharp D.C., Nilson J.H.;
RT "A single gene encodes the beta-subunits of equine luteinizing
RT hormone and chorionic gonadotropin."
RL Mol. Endocrinol. 6:951-959(1992).
RN [2]
RP SEQUENCE OF 21-169.
RX MEDLINE; 87250476.
RA Bousfield G.R., Liu W.-K., Sugino H., Ward D.N.;
RT "Structural studies on equine glycoprotein hormones. Amino acid
RT sequence of equine lutropin beta-subunit."
RL J. Biol. Chem. 262:8610-8620(1987).
RN [3]
RP SEQUENCE OF 21-169.

RX MEDLINE; 87250475.
RA Sugino H., Bousfield G.R., Moore W.T. Jr., Ward D.N.;
RT "Structural studies on equine glycoprotein hormones. Amino acid
RT sequence of equine chorionic gonadotropin beta-subunit."
RL J. Biol. Chem. 262:8609-8609(1987).
RN [4]
RP CHORIOGONADOTROPIN, PARTIAL SEQUENCE.
RA Ward D.N., Moore W.T. Jr., Burleigh B.D.;
RT "Structural studies on equine chorionic gonadotropin."
RL J. Protein Chem. 1:263-280(1982).
RN [5]
RP STRUCTURE OF CARBOHYDRATES.
RX MEDLINE; 90233854.
RA Damm J.B.L., Hard K., Kamerling J.P., van Dedem G.W.K.,
RA Vlieghehart J.F.G.;
RT "Structure determination of the major N- and O-linked carbohydrate
RT chains of the beta subunit from equine chorionic gonadotropin."
RL Eur. J. Biochem. 189:175-183(1990).
CC -!- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
CC THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- PTM: AT LEAST FOUR O-GLYCOSYLATION SITES ARE PRESENT.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S41704; AAB22775.1; -.
CC PIR; A01503; KTHOB.
CC PIR; A29304; A29304.
CC PIR; A29305; A29305.
CC PIR; A41917; A41917.
CC HSP; P01233; IHRP.
CC PRINTS; PR00438; GFCYSKNOT.
CC PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
CC PFAM; PF00007; Cys_knot; 1.
CC Hormone; Glycoprotein; Signal.
CC SIGNAL 1 20 LUTROPIN/CHORIOGONADOTROPIN BETA CHAIN.
CC CHAIN 21 169 BY SIMILARITY.
CC DISULFID 29 77 BY SIMILARITY.
CC DISULFID 43 92 BY SIMILARITY.
CC DISULFID 46 130 BY SIMILARITY.
CC DISULFID 54 108 BY SIMILARITY.
CC DISULFID 58 110 BY SIMILARITY.
CC DISULFID 113 120 BY SIMILARITY.
CC CARBOHYD 33 33 GLYCAN SHOWING A REMARKABLY STRUCTURAL
CC HETEROGENEITY.
CC SEQUENCE 169 AA; 17865 MW; 1244ADBEB843EF1A CRC64;

Query Match 52.6%; Score 90; DB 1; Length 169;
Best Local Similarity 50.0%; Pred. No. 4e-06;
Matches 15; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgcpbhvafsyypva 30
||||| : : : ||||| | | |||||
Db 76 VCTYRELREASIRLPGCPGVDPWVSFPVA 105

RESULT 36
LSHB_BALAC
ID LSHB_BALAC STANDARD; PRT; 118 AA.
AC F33088;
DT 01-OCT-1993 (Rel. 27, Created)
```


DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DE LUTROPIN BETA CHAIN (LUTEINIZING HORMONE) (LSH-B) (LH-B).
 GN LHB.
 OS Balaenoptera acutorostrata (Minke whale) (Lesser rorqual).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Cetartiodactyla; Cetacea; Mysticeti; Balaenopteridae;
 OC Balaenoptera.
 RN [1]
 RP SEQUENCE.
 RA Karasev V.S., Pankov Y.A.;
 RT "Amino acid sequence of reduced and carboxymethylated alpha- and beta-
 subunits of the little picked whale luteinizing hormone.";
 RL Biokhimia 50:1972-1986(1985).
 CC -!- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
 THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
 CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
 CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
 LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
 FAMILY.
 CC PIR; P0139; PNO139.
 DR HSSP; P01233; 1HRP.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
 DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; FALSE_NEG.
 DR PFAM; PF00007; Cys_knot; 1.
 KW Hormone; Glycoprotein.
 FT DISULFID 9 57 BY SIMILARITY.
 FT DISULFID 23 72 BY SIMILARITY.
 FT DISULFID 26 110 BY SIMILARITY.
 FT DISULFID 34 88 BY SIMILARITY.
 FT DISULFID 38 90 BY SIMILARITY.
 FT DISULFID 93 100 BY SIMILARITY.
 FT CARBOHYD 13 13
 SQ SEQUENCE 118 AA; 12414 MW; 039F229EFC480F5D CRC64;

Query Match 52.0%; Score 89; DB 1; Length 118;
 Best Local Similarity 50.0%; Pred. No. 3.8e-06;
 Matches 15; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgclhvpafsyypva 30
 ||||| : : : : : ||||| : : : : :
 Db 56 VCTYRLRFASIRLPGCPGVPBMVSPVA 85

RESULT 37
 LSHB_EQUAS STANDARD; PRT; 169 AA.
 ID LSHB_EQUAS
 AC P19794;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE LUTROPIN/CHORIOGONADOTROPIN BETA CHAIN PRECURSOR (LSH-B/CG-B)
 DE (LUTEINIZING HORMONE BETA SUBUNIT).
 GN LHB.
 OS Equus asinus (Donkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Perissodactyla; Equidae; Equus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chopineau M., Combarnous Y., Allen W.R., Stewart F.;
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP PRELIMINARY SEQUENCE OF 105-169 FROM N.A.
 RC TISSUE=PLACENTA;
 RX MEDLINE; 90262634.
 RA Leigh S.E.A., Stewart F.;
 RT "Partial cDNA sequence for the donkey chorionic gonadotrophin-beta
 subunit suggests evolution from an ancestral LH-beta gene.";
 RL J. Mol. Endocrinol. 4:143-150(1990).
 CC -!- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING

CC THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
 CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
 CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
 LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X80116; CAA58422.1; -; -;
 DR EMBL; X53669; CAA37709.1; ALT_SEQ.
 DR PIR; S15676; S15676.
 DR HSSP; P01233; 1HRP.
 DR PRINTS; PR00438; GFCYSKNOT.
 DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
 DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
 DR PFAM; PF00007; Cys_knot; 1.
 KW Hormone; Glycoprotein; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 169 LUTROPIN/CHORIOGONADOTROPIN BETA CHAIN.
 FT DISULFID 29 77 BY SIMILARITY.
 FT DISULFID 43 92 BY SIMILARITY.
 FT DISULFID 46 130 BY SIMILARITY.
 FT DISULFID 54 108 BY SIMILARITY.
 FT DISULFID 58 110 BY SIMILARITY.
 FT DISULFID 113 120 BY SIMILARITY.
 FT CARBOHYD 33 33 PROBABLE.
 SQ SEQUENCE 169 AA; 17943 MW; AEED20891E96FA7C CRC64;

Query Match 52.0%; Score 89; DB 1; Length 169;
 Best Local Similarity 50.0%; Pred. No. 5.6e-06;
 Matches 15; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Qy 1 vctyrdfiyrtveipgclhvpafsyypva 30
 ||||| : : : : : ||||| : : : : :
 Db 76 VCTYRELFGSIRLPGCPGVPBMVSPVA 105

RESULT 38
 LSHB_PHYCA STANDARD; PRT; 118 AA.
 ID LSHB_PHYCA
 AC P25330;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE LUTROPIN BETA CHAIN (LUTEINIZING HORMONE) (LSH-B) (LH-B).
 GN LHB.
 OS Physeter catodon (Sperm whale) (Physeter macrocephalus).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Physeteridae;
 OC Physeter.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE; 87032654.
 RA Pankov Y.A., Karasev V.S.;
 RT "Primary structure of sperm whale luteinizing hormone.";
 RL Int. J. Pept. Protein Res. 28:124-129(1986).
 RN [2]
 RP SEQUENCE.
 RX MEDLINE; 84281133.
 RA Pankov Y.A., Karasev V.S.;
 RT "Luteinizing hormone of the sperm whale. Amino acid sequences of
 reduced and carboxymethylated beta-subunits.";
 RL Biokhimia 49:1004-1018(1984).
 CC -!- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
 THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.

CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN. HORMONES BETA CHAIN
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC PIR; PN0141; DR
CC HSP; P01233; 1HRP.
CC PRINTS; PR00438; GFCYSKNOT.
CC PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
CC PFAM; PF00007; Cys_knot; 1.
CC Hormone; Glycoprotein.
KW FT DISULFID 9 57 BY SIMILARITY.
FT DISULFID 23 72 BY SIMILARITY.
FT DISULFID 26 110 BY SIMILARITY.
FT DISULFID 34 88 BY SIMILARITY.
FT DISULFID 38 90 BY SIMILARITY.
FT DISULFID 93 100 BY SIMILARITY.
FT CARBOHYD 13 13
SQ SEQUENCE 118 AA; 12412 MW; 81177A56382F15E7 CRC64;

Query Match 51.5%; Score 88; DB 1; Length 118;
Best Local Similarity 50.9%; Pred. No. 5.3e-06;
Matches 15; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgclhvpafsyypva 30
||||| : : : ||||| | | : |||||

Db 56 VCTYRQLRFASIRLPGCPGVNPMVSPFA 85

RESULT 39
GTHB_MURCI STANDARD; PRT; 113 AA.
AC P12837;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GONADOTROPIN BETA CHAIN.
OS Muranesox cinereus (Pike eel) (Dagtertooth pike conger).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
CC Neopterygii; Teleostei; Anguilliformes; Muraenesocidae; Muranesox.
RN [1]
RP SEQUENCE.
RC TISSUE-PITUITARY;
RX MEDLINE; 90092087.
RA Liu C.-S., Huang F.-L., Chang Y.-S., Lo T.-B.;
RT "Pike eel (Muranesox cinereus) gonadotropin. Amino acid sequences of
both alpha and beta subunits.";
RL Eur. J. Biochem. 186:105-114(1989).
CC -1- FUNCTION: INVOLVED IN GAMETOGENESIS AND STEROIDGENESIS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC PIR; S07092; S07092.
CC HSP; P01233; 1HRP.
CC PRINTS; PR00438; GFCYSKNOT.
CC PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
CC PFAM; PF00007; Cys_knot; 1.
CC Hormone; Glycoprotein.
KW FT DISULFID 6 54 BY SIMILARITY.
FT DISULFID 20 69 BY SIMILARITY.
FT DISULFID 23 107 BY SIMILARITY.
FT DISULFID 31 85 BY SIMILARITY.
FT DISULFID 35 87 BY SIMILARITY.
FT DISULFID 90 97 BY SIMILARITY.
FT CARBOHYD 10 10
SQ SEQUENCE 113 AA; 12582 MW; 17DD0ADF6ACA6F7 CRC64;

Query Match 50.9%; Score 87; DB 1; Length 113;
Best Local Similarity 53.3%; Pred. No. 7.1e-06;

Matches 16; Conservative 4; Mismatches 10; Indels 0; Gaps 0;
QY 1 vctyrdfiyrtveipgclhvpafsyypva 30
||||| | | | | | | | | | : |||||
Db 53 VCTYRDVRYETVRLPDCRPGVDPHVTFPVA 82
RESULT 40
GTH2_ANGAN STANDARD; PRT; 140 AA.
ID GTH2_ANGAN
AC P27767;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GONADOTROPIN BETA-II CHAIN PRECURSOR (GTH-II).
OS Anguilla anguilla (European freshwater eel).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
CC Neopterygii; Teleostei; Anguilliformes; Anguillidae; Anguilla.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PITUITARY;
RX MEDLINE; 90334705.
RA Querat B., Mounni M., Jutisz M., Fontaine Y.A., Counis R.;
RT "Molecular cloning and sequence analysis of the cDNA for the putative
beta subunit of the type-II gonadotropin from the European eel.";
RL J. Mol. Endocrinol. 4:257-264(1990).
CC -1- FUNCTION: INVOLVED IN GAMETOGENESIS AND STEROIDGENESIS.
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X61039; CAA43374.1; .
CC PIR; A48166; A48166.
CC HSP; P01233; 1HRP.
CC PRINTS; PR00438; GFCYSKNOT.
CC PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
CC PFAM; PF00007; Cys_knot; 1.
CC Signal; Hormone; Glycoprotein.
FT SIGNAL 1 24 BY SIMILARITY.
FT CHAIN 25 140 GONADOTROPIN BETA-II CHAIN.
FT DISULFID 30 78 BY SIMILARITY.
FT DISULFID 44 93 BY SIMILARITY.
FT DISULFID 47 131 BY SIMILARITY.
FT DISULFID 55 109 BY SIMILARITY.
FT DISULFID 59 111 BY SIMILARITY.
FT DISULFID 114 121 BY SIMILARITY.
FT CARBOHYD 34 34 POTENTIAL.
SQ SEQUENCE 140 AA; 15478 MW; 10312C061717DD80 CRC64;

Query Match 50.9%; Score 87; DB 1; Length 140;
Best Local Similarity 53.3%; Pred. No. 8.9e-06;
Matches 16; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgclhvpafsyypva 30
||||| | | | | | | | | | : |||||

Db 77 VCTYRDVRYETVRLPDCRPGVDPHVTFPVA 106

RESULT 41
GTH2_CORAU STANDARD; PRT; 142 AA.
ID GTH2_CORAU
AC P48251;
DT 01-FEB-1996 (Rel. 33, Created)


```
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GONADOTROPIN BETA-II CHAIN PRECURSOR (GTH-II).
OS Coregonus autumnalis (Baikal omul).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
OC Salmoniformes; Salmonidae; Coregonus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PITUITARY;
RX MEDLINE; 95082790.
RA Trofimova I.N., Belikov S.I.;
RT "Cloning and sequencing the cDNA for the beta-subunit of Baikal omul
gonadotropin.";
RL Mol. Biol. (Mosk) 28:1052-1056(1994).
CC -!- FUNCTION: INVOLVED IN GAMETOGENESIS AND STEROIDOGENESIS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L23431; AAA68207.1; -.
DR HSP; P01233; 1HRP.
DR PRINTS; PR00438; GFCYSKNOT.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
DR PFAM; PF00007; Cys_knot; 1.
KW Hormone; Glycoprotein; Signal.
FT SIGNAL 1 24
FT CHAIN 25 142 GONADOTROPIN BETA-II CHAIN.
FT DISULFID 30 78 BY SIMILARITY.
FT DISULFID 44 93 BY SIMILARITY.
FT DISULFID 47 131 BY SIMILARITY.
FT DISULFID 55 109 BY SIMILARITY.
FT DISULFID 59 111 BY SIMILARITY.
FT DISULFID 114 121 BY SIMILARITY.
FT CARBOHYD 34 34 POTENTIAL.
SQ SEQUENCE 142 AA; 15843 MW; 21105B70B410797D CRC64;

Query Match 50.9%; Score 87; DB 1; Length 142;
Best Local Similarity 53.3%; Pred. No. 9.1e-06;
Matches 16; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 1 vctyrdfyrtveipgcpvhvafysypva 30
||||| | : | | | | : | | | |
Db 77 VCTYRDVRYETICLPDCSPWDPHVTYPVA 106

RESULT 42
GTHB_ONCTS
ID GTHB_ONCTS STANDARD; PRT; 142 AA.
AC P07732;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GONADOTROPIN BETA CHAIN PRECURSOR (BETA-GTH).
OS Oncorhynchus tshawytscha (Chinook salmon) (King salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
OC Salmoniformes; Salmonidae; Oncorhynchus.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE; 87004682.
RA Trinh K.-Y., Wang N.C., Hew C.-L., Crim L.W.;
RT "Molecular cloning and sequencing of salmon gonadotropin beta
```

```
subunit.";
RT Eur. J. Biochem. 159:619-624(1986).
CC -!- FUNCTION: INVOLVED IN GAMETOGENESIS AND STEROIDOGENESIS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04404; CAA27992.1; -.
DR PIR; A25800; A25800.
DR HSP; P01233; 1HRP.
DR PRINTS; PR00438; GFCYSKNOT.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
DR PFAM; PF00007; Cys_knot; 1.
KW Hormone; Glycoprotein; Signal.
FT SIGNAL 1 23
FT CHAIN 24 142 GONADOTROPIN BETA CHAIN.
FT DISULFID 29 77 BY SIMILARITY.
FT DISULFID 43 92 BY SIMILARITY.
FT DISULFID 46 130 BY SIMILARITY.
FT DISULFID 54 108 BY SIMILARITY.
FT DISULFID 58 110 BY SIMILARITY.
FT DISULFID 113 120 BY SIMILARITY.
FT CARBOHYD 33 33 PROBABLE.
SQ SEQUENCE 142 AA; 15902 MW; BEAE8EBFFD1292F8 CRC64;

Query Match 50.9%; Score 87; DB 1; Length 142;
Best Local Similarity 50.0%; Pred. No. 9.1e-06;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 1 vctyrdfyrtveipgcpvhvafysypva 30
||||| | : | | | | : | | | |
Db 76 VCTYRDVRYEMIRLPDCPPWDPHVTYPVA 105

RESULT 43
LSHB_BOVIN
ID LSHB_BOVIN STANDARD; PRT; 141 AA.
AC P04651;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE) (LSH-B) (LH-B).
GN LHB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 85207729.
RA Virgin J.B., Silver B.J., Thomason A.R., Nilson J.H.;
RT "The gene for the beta subunit of bovine luteinizing hormone encodes
a gonadotropin mRNA with an unusually short 5'-untranslated region.";
RL J. Biol. Chem. 260:7072-7077(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 85182575.
RA Maurer R.A.;
RT "Analysis of several bovine lutropin beta subunit cDNAs reveals
heterogeneity in nucleotide sequence.";
RL J. Biol. Chem. 260:4684-4687(1985).
RN [3]
RP SEQUENCE OF 21-139.
```



```

FT DISULFID 58 110 BY SIMILARITY.
FT DISULFID 113 120
FT MOD_RES 21 21 BLOCKED.
FT CARBOHYD 33 33
FT VARIANT 138 141 MISSING (IN SOME MOLECULES).
FT CONFLICT 30 30 Q -> E (IN REF. 4).
FT CONFLICT 59 59 L -> P (IN REF. 1).
FT CONFLICT 63 63 R -> Q (IN REF. 2).
FT CONFLICT 71 72 PW -> PEM (IN REF. 4).
FT CONFLICT 81 81 E -> Q (IN REF. 4).
FT CONFLICT 122 123 GP -> PG (IN REF. 3 AND 4).
FT CONFLICT 126 126 Q -> E (IN REF. 3 AND 4).
SQ SEQUENCE 141 AA; 15184 MW; C59EC7C0AA55A9DC CRC64;

Query Match 50.3%; Score 86; DB 1; Length 141;
Best Local Similarity 50.0%; Pred. NO. 1.3e-05;
Matches 15; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Oy 1 vctyrdfiyrtveipgcpvhvapyfsypva 30
Db 76 VCTYHRLFAFASVRLPGCPGVDPMVSFVPA 105

RESULT 45
CGHB_HUMAN
ID CGHB_HUMAN STANDARD; PRT; 165 AA.
AC P01233; Q14000; Q13991;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-FEB-2000 (Rel. 39, Last annotation update)
DE CHORIONOTROPIN BETA CHAIN PRECURSOR (CHORIONIC GONADOTROPIN BETA
DE SUBUNIT).
GN CGB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RX MEDLINE; 81012134.
RA Fiddes J.C., Goodman H.M.;
RT "The cDNA for the beta-subunit of human chorionic gonadotropin
RT suggests evolution of a gene by readthrough into the 3'-untranslated
RT region.";
RL Nature 286:684-687(1980).
RN [2]
RX MEDLINE; 84093590.
RA Talmadge K., Vamvakopoulos N.C., Fiddes J.C.;
RT "Evolution of the genes for the beta subunits of human chorionic
RT gonadotropin and luteinizing hormone.";
RL Nature 307:37-40(1984).
RN [3]
RX MEDLINE; 84008141.
RA Policastro P., Ovitt C.E., Hoshina M., Fukuoaka H., Boothby M.R.,
RA Boime I.;
RT "The beta subunit of human chorionic gonadotropin is encoded by
RT multiple genes.";
RL J. Biol. Chem. 258:11492-11499(1983).
RN [4]
RX MEDLINE; 81117268.
RA Birken S., Fetherston J., Canfield R.E., Boime I.;
RT "The amino acid sequences of the prepeptides contained in the alpha
RT and beta subunits of human chorionic gonadotropin.";
RL J. Biol. Chem. 256:1816-1823(1981).
RN [5]
RX MEDLINE; 75211304.
RA Morgan F.J., Birken S., Canfield R.E.;
RT "The amino acid sequence of human chorionic gonadotropin. The alpha
RT subunit and beta subunit.";
```

```

J. Biol. Chem. 250:5247-5258(1975).
RN [6]
RX PRELIMINARY SEQUENCE OF 21-165.
MEDLINE; 74011267.
RA Carlsen R.B., Bahl O.P., Swaminathan N.;
RT "Human chorionic gonadotropin. Linear amino acid sequence of the beta
RT subunit.";
RL J. Biol. Chem. 248:6810-6827(1973).
RN [7]
RX SEQUENCE OF 1-5 FROM N.A.
MEDLINE; 86195987.
RA Policastro P.F., Daniels-McQueen S., Carle G., Boime I.;
RT "A map of the hCG beta-LH beta gene cluster.";
RL J. Biol. Chem. 261:5907-5916(1986).
RN [8]
RX PRELIMINARY ASSIGNMENT OF DISULFIDE BONDS.
MEDLINE; 81215630.
RA Mise T., Bahl O.P.;
RT "Assignment of disulfide bonds in the beta subunit of human chorionic
RT gonadotropin.";
RL J. Biol. Chem. 256:6587-6592(1981).
RN [9]
RX DISULFIDE BONDS.
MEDLINE; 90094415.
RA Saccuzo Beebe J., Mountjoy K., Krzesicki R.F., Perini F.,
RA Rudden R.W.;
RT "Role of disulfide bond formation in the folding of human chorionic
RT gonadotropin beta subunit into an alpha beta dimer assembly-competent
RT form.";
RL J. Biol. Chem. 265:312-317(1990).
RN [10]
RX STRUCTURE OF CARBOHYDRATES.
MEDLINE; 92314469.
RA Weisshaar G., Hiya J., Renwick A.G.C.;
RT "Site-specific N-glycosylation of human chorionic gonadotropin --
RT structural analysis of glycopeptides by one- and two-dimensional 1H
RT NMR spectroscopy";
RL Glycobiology 1:393-404(1991).
RN [11]
RX X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
MEDLINE; 94261179.
RA Laphorn A.J., Harris D.C., Littlejohn A., Lustbader J.W.,
RA Canfield R.E., Machin K.J., Morgan F.J., Isaacs N.W.;
RT "Crystal structure of human chorionic gonadotropin.";
RL Nature 369:455-461(1994).
CC -!- FUNCTION: STIMULATES THE OVARIAN TO SYNTHESIZE THE STEROIDS THAT
CC ARE ESSENTIAL FOR THE MAINTENANCE OF PREGNANCY.
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- TISSUE SPECIFICITY: PLACENTA.
CC -!- DEVELOPMENTAL STAGE: MADE BY THE FIRST TRIMESTER PLACENTA.
CC -!- PHARMACEUTICAL: AVAILABLE UNDER THE NAMES NOVAREL (FERRING) AND
CC PROFASI (SERONO).
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J00117; AAA96690.1; -
CC EMBL; X00266; CAA25069.1; ALT_INIT.
CC EMBL; M13504; AAA52005.1; -
CC EMBL; M13505; AAA52008.1; -
CC EMBL; M13503; AAA52009.1; -
CC EMBL; K03189; AAA53288.1; -
CC EMBL; K03187; AAA53288.1; JOINED.
CC EMBL; K03188; AAA53288.1; JOINED.
```


DR EMBL; K03183; AAA53287.1; .
DR EMBL; K00092; AAA53287.1; JOINED.
DR EMBL; K03182; AAA53287.1; JOINED.
DR PIR; A01502; KTHUB.
DR PDB; 1HCN; 30-SEP-94.
DR PDB; 1HRP; 01-NOV-94.
DR PDB; 1XUL; 15-MAY-97.
DR MIM; 118860; .
DR PRINTS; PR00438; GFCYSKNOT.

DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
DR PFAM; PF00007; Cys_knot; 1.
KW Hormone; Glycoprotein; Signal; Pharmaceutical; 3D-structure.

FT SIGNAL 1 20
FT CHAIN 21 165 CHORIOGONADOTROPIN BETA CHAIN.

FT DISULFID 29 77
FT DISULFID 43 92
FT DISULFID 46 130
FT DISULFID 54 108
FT DISULFID 58 110
FT DISULFID 113 120
FT CARBOHYD 33 33
FT CARBOHYD 50 50
FT CARBOHYD 141 141
FT CARBOHYD 147 147
FT CARBOHYD 152 152
FT CARBOHYD 158 158
FT VARIANT 137 137
D -> A (IN GENE 6).
/FTID=VAR_003188.

FT CONFLICT 24 24
FT SEQUENCE 165 AA; 17739 MW; 5598FB9E51A05748 CRC64;

Query Match 50.3%; Score 86; DB 1; Length 165;
Best Local Similarity 50.0%; Pred. No. 1.5e-05;
Matches 15; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 vctyrdfyrtveipgpcphvafsyypva 30
|| ||| : : : |||| | | |||

DB 76 VCNRYDRVFESIRLPGCRGNPVWSYAVA 105

RESULT 46

LSHB_RANCA STANDARD; PRT; 112 AA.

AC P80071;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE LUTROPIN BETA CHAIN (LUTEINIZING HORMONE) (LSH-B) (LH-B).
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Anura; Neobatrachia; Ranolia; Ranidae; Raninae; Rana.
RN [1]
RP SEQUENCE.

RA MEDLINE; 92209490.
RX Hiroaki H., Tomoko H., Yoichi H.;
RT "Amphibian lutropin from the bullfrog Rana catesbeiana. Complete amino acid sequence of the beta subunit."
RL Eur. J. Biochem. 205:105-110(1992).
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN, LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN FAMILY.

CC PIR; S21196; S21196.
CC HSP; P01233; IHRP.
CC PRINTS; PR00438; GFCYSKNOT.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
DR PFAM; PF00007; Cys_knot; 1.
KW Hormone; Glycoprotein.
FT DISULFID 4 52 BY SIMILARITY.

FT DISULFID 18 67 BY SIMILARITY.
FT DISULFID 21 105 BY SIMILARITY.
FT DISULFID 29 83 BY SIMILARITY.
FT DISULFID 33 85 BY SIMILARITY.
FT DISULFID 88 95 BY SIMILARITY.
FT CARBOHYD 8 8
SQ SEQUENCE 112 AA; 12675 MW; 3DF55E5CD91D1071 CRC64;

Query Match 49.7%; Score 85; DB 1; Length 112;
Best Local Similarity 43.3%; Pred. No. 1.4e-05;
Matches 13; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 vctyrdfyrtveipgpcphvafsyypva 30
: ||| : : : |||| | | |||

DB 51 ICTYKEIRYDTIKLPDCLPGTDPFFTPVA 80

RESULT 47

LSHB_CANFA STANDARD; PRT; 138 AA.

AC P18842;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE) (LSH-B) (LH-B) (FRAGMENT).
DE LHB.
GN LHB.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88096605.
RX Wolf D.L., Appleby V.I., Hjerrild K., Baker A.R., Talmadge K.;
RT "Nucleic acid and amino acid sequences of dog beta LH: comparison to rat, cow and human beta LH".
RL Nucleic Acids Res. 15:10602-10602(1987).
CC -!- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN, LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN FAMILY.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y00518; CAA68572.1; .
CC PIR; S00512; S00512.
CC HSP; P01233; IHRP.
CC DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
CC DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
CC DR PFAM; PF00007; Cys_knot; 1.
CC KW Hormone; Signal; Glycoprotein.
CC FT NON_TER 1 1
CC FT SIGNAL <1 17 LUTROPIN BETA CHAIN.
CC FT CHAIN 18 138 BY SIMILARITY.
CC FT DISULFID 26 74 BY SIMILARITY.
CC FT DISULFID 40 89 BY SIMILARITY.
CC FT DISULFID 43 127 BY SIMILARITY.
CC FT DISULFID 51 105 BY SIMILARITY.
CC FT DISULFID 55 107 BY SIMILARITY.
CC FT DISULFID 110 117 BY SIMILARITY.
CC FT CARBOHYD 30 30 PROBABLE.
CC SEQUENCE 138 AA; 14594 MW; E3639FE6B03F1948 CRC64;

Query Match 49.7%; Score 85; DB 1; Length 138;
Best Local Similarity 46.7%; Pred. No. 1.7e-05;
Matches 14; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpplhvafsyypva 30
|||| : : : ||||| | | : ||||
Db 73 VCIYHELFAFASIRLPGCPGVDPMVSPVA 102

RESULT 48

ID GTH2_FUNHE STANDARD; PRT; 136 AA.
AC P30972;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GONADOTROPIN BETA-II CHAIN PRECURSOR (GTH-II).
OS Fundulus heteroclitus (Killifish) (Mummichog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Atherinomorpha;
OC Cyprinodontiformes; Cyprinodontoidae; Fundulidae; Fundulus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PITUITARY;
RX MEDLINE; 92405806.
RA Lin Y.-W.P., Rupnow B.A., Price D.A., Greenberg R.M., Wallace R.A.;
RT "Fundulus heteroclitus gonadotropins. 3. Cloning and sequencing of
RT gonadotropic hormone (GTH) I and II beta-subunits using the
RT polymerase chain reaction.";
RL Mol. Cell. Endocrinol. 85:127-139(1992).
CC -|- FUNCTION: INVOLVED IN GAMETOGENESIS AND STEROIDGENESIS.
CC -|- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -|- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M87015; AAB59963.1; -;
CC HSP; P01233; 1HRP.
CC PRINTS; PR00438; GFCYSKNOT.
CC PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
CC PFAM; PF00007; Cys_knot; 1.
CC Signal; Hormone; Glycoprotein.
CC SIGNAL 1 21 POTENTIAL.
CC CHAIN 22 136 GONADOTROPIN BETA-II CHAIN.
FT DISULFID 27 75
FT DISULFID 41 90 BY SIMILARITY.
FT DISULFID 44 128 BY SIMILARITY.
FT DISULFID 52 106 BY SIMILARITY.
FT DISULFID 56 108 BY SIMILARITY.
FT DISULFID 111 118 BY SIMILARITY.
FT CARBOHYD 31 31 POTENTIAL.
SQ SEQUENCE 136 AA; 15016 MW; 716576844A6A1653 CRC64;

Query Match 48.5%; Score 83; DB 1; Length 136;
Best Local Similarity 53.3%; Pred. No. 3.3e-05;
Matches 16; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpplhvafsyypva 30
|||| : : : ||||| | | : ||||
Db 74 VCTYGLDYIKTFFPEFCVPGVDPMVTPVA 103

RESULT 49

CGHB_PAPAN
ID CGHB_PAPAN STANDARD; PRT; 165 AA.
AC P07434;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CHORIOGONADOTROPIN BETA CHAIN PRECURSOR (CHORIONIC GONADOTROPIN BETA
DE SUBUNIT).
GN CGB.
OS Papio anubis (Olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Cercopitheciidae; Cercopitheciinae;
OC Papio.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 87106851.
RA Crawford R.J., Tregear G.W., Niall H.D.;
RT "The nucleotide sequences of baboon chorionic gonadotropin
RT beta-subunit genes have diverged from the human.";
RL Gene 46:161-169(1986)
CC -|- FUNCTION: STIMULATES THE OVARIES TO SYNTHESIZE THE STEROIDS THAT
CC ARE ESSENTIAL FOR THE MAINTENANCE OF PREGNANCY.
CC -|- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -|- TISSUE SPECIFICITY: PLACENTA.
CC -|- MISCELLANEOUS: THERE ARE AT LEAST FIVE COPIES OF CG-RELATED GENES
CC AND AT LEAST TWO OF THESE ARE EXPRESSED IN THE BABOON PLACENTA.
CC -|- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M14966; AAA35383.1; -;
CC PIR; A25808; KTBAB.
CC HSP; P01233; 1HRP.
CC PRINTS; PR00438; GFCYSKNOT.
CC PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
CC PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
CC PFAM; PF00007; Cys_knot; 1.
CC Hormone; Glycoprotein; Signal.
CC SIGNAL 1 20 BY SIMILARITY.
CC CHAIN 21 165 CHORIOGONADOTROPIN BETA CHAIN.
FT DISULFID 29 77 BY SIMILARITY.
FT DISULFID 43 92 BY SIMILARITY.
FT DISULFID 46 130 BY SIMILARITY.
FT DISULFID 54 108 BY SIMILARITY.
FT DISULFID 58 110 BY SIMILARITY.
FT DISULFID 113 120 BY SIMILARITY.
FT DISULFID 33 33 BY SIMILARITY.
FT CARBOHYD 50 50 BY SIMILARITY.
FT CARBOHYD 141 141 BY SIMILARITY.
FT CARBOHYD 147 147 BY SIMILARITY.
FT CARBOHYD 152 152 BY SIMILARITY.
SQ SEQUENCE 165 AA; 17592 MW; 36D3E207A9F1E1C3 CRC64;

Query Match 47.4%; Score 81; DB 1; Length 165;
Best Local Similarity 46.7%; Pred. No. 7.8e-05;
Matches 14; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpplhvafsyypva 30
|||| : : : ||||| | | : ||||
Db 76 VCNTRYEVRESIRLPGCPGVDPMVSPVA 105

RESULT 50


```
GTH1_ONCKE
ID  GTH1_ONCKE  STANDARD;  PRT;  137 AA.
AC  P10257;
DT  01-MAR-1989 (Rel. 10, Created)
DT  01-JAN-1990 (Rel. 13, Last sequence update)
DT  15-JUL-1999 (Rel. 38, Last annotation update)
DE  GONADOTROPIN BETA-I CHAIN PRECURSOR (GTH-I).
OS  Oncorhynchus keta (Chum salmon).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC  Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
OC  Salmoniformes; Salmonidae; Oncorhynchus.
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE; 90046849.
RA  Sekine S., Saito A., Itoh H., Kawauchi H., Itoh S.;
RT  "Molecular cloning and sequence analysis of chum salmon gonadotropin
RL  CDNAS.";
RL  Proc. Natl. Acad. Sci. U.S.A. 86:8645-8649(1989).
RN  [2]
RP  SEQUENCE OF 25-137.
RX  MEDLINE; 89053031.
RA  Itoh H., Suzuki K., Kawauchi H.;
RT  "The complete amino acid sequences of beta-subunits of two distinct
RL  chum salmon GTHs.";
RL  Gen. Comp. Endocrinol. 71:438-451(1988).
CC  -!- FUNCTION: INVOLVED IN GAMETOGENESIS AND STEROIDOGENESIS.
CC  -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC  -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC  FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M27153; AAA49408.1; -
DR  FIR; B36179; B36179.
DR  HSP; P01233; IHRP.
DR  PRINTS; PR00438; GFCYSKNOT.
DR  PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR  PROSITE; PS00689; GLYCO_HORMONE_BETA_3; 1.
DR  PFAM; PF00007; Cys_knot; 1.
KW  Hormone; Glycoprotein; Signal.
FT  SIGNAL 1 24
FT  CHAIN 25 137 GONADOTROPIN BETA-I CHAIN.
FT  DISULFID 32 78 BY SIMILARITY.
FT  DISULFID 46 93 BY SIMILARITY.
FT  DISULFID 55 108 BY SIMILARITY.
FT  DISULFID 59 110 BY SIMILARITY.
FT  DISULFID 113 120 BY SIMILARITY.
FT  CARBOHYD 36 36
FT  CONFLICT 67 67 Q -> E (IN REF. 2).
SQ  SEQUENCE 137 AA; 15442 MW; B040BD6AF2A7DB3B CRC64;

Query Match 40.1%; Score 68.5; DB 1; Length 137;
Best Local Similarity 43.3%; Pred. No. 0.0041;
Matches 13; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 1 vctyrdifvtrveipgcpnhvayfysypva 30
|| : : : | : | | | | | | | |
DB 77 VCNFKENSYEKVYLEGCPSCGVEPFF-IPVA 105
```

Search completed: July 14, 2000, 09:33:58
Job time: 2289 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 14, 2000, 08:55:49 ; Search time 50.59 Seconds

(without alignments)
41.115 Million cell updates/sec

Title: BETA-CHAIN

Perfect score: 171

Sequence: 1 vctyrdfyrtveipgcpvhvapyfsypva 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 225878 seqs, 69334122 residues

Total number of hits satisfying chosen parameters: 225877

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 99%

Listing first 1000 summaries

Database :

1: sp_archea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	164	95.9	87	4	Q15962 homo sapien
2	107	62.6	150	13	O13051 cyprinus ca
3	106	62.0	137	13	Q90225 acanthopagr
4	101	59.1	140	13	Q98849 carassius a
5	101	59.1	147	13	O13052 ctenopharyn
6	99	57.9	147	13	Q91121 morone saxa
7	98	57.3	141	6	O46482 trichosurus
8	97	56.7	150	13	O13049 carassius a
9	95	55.6	127	13	Q9YGG3 anguilla ja
10	93	54.4	141	11	O60844 mus musculus
11	91	53.2	139	11	O62778 rattus norv
12	91	53.2	143	6	O77805 felis silve
13	90	52.6	88	13	Q9W6Q7 salmo salar
14	90	52.6	89	6	O46618 equus zebra
15	90	52.6	89	6	O46619 equus hemio
16	85	49.7	82	6	O46622 ceratotheri
17	85	49.7	135	6	O19102 ceratotheri
18	85	49.7	141	6	O77835 ceratotheri
19	83	48.5	130	13	Q9YIB3 carassius a

20	82	48.0	149	13	Q9YGH2
21	80.5	47.1	80	11	O63013 rattus norv
22	80	46.8	130	13	O13050 cyprinus ca
23	80	46.8	130	13	Q98848 carassius a
24	67.5	39.5	123	13	Q9W6Q8
25	56	32.7	262	2	O67006
26	49	28.7	108	12	O66148 cocksfoot m
27	49	28.7	120	13	Q91120 morone saxa
28	49	28.7	292	10	O48730 arabidopsis
29	49	28.7	452	2	P94612 coxiella bu
30	49	28.7	813	11	O63315 rattus norv
31	49	28.7	813	11	Q9WTP5 mus musculus
32	48.5	28.4	47	2	O50041 mycobacteri
33	48.5	28.4	491	5	O18351 caenorhabdi
34	48	28.1	148	1	O26497 methanobact
35	48	28.1	307	2	O49138 methylobact
36	48	28.1	327	2	Q9ZIS9 escherichia
37	47.5	27.8	374	3	Q93289 saccharomyc
38	47	27.5	320	2	O68306 nostoc pcc7
39	47	27.5	359	2	O88036 streptomyc
40	47	27.5	360	2	O53282 mycobacteri
41	47	27.5	383	5	O01838 caenorhabdi
42	47	27.5	448	2	O67030 aquifex aeo
43	47	27.5	1645	5	O27448 b glutamine
44	46.5	27.2	289	5	O07338 pyemotes tr
45	46.5	27.2	334	5	Q9XUN4 caenorhabdi
46	46	26.9	341	10	O49857 glycine max
47	46	26.9	341	10	Q39856 glycine max
48	46	26.9	380	4	O43205 homo sapien
49	46	26.9	382	3	O74497 schizosacch
50	46	26.9	711	2	O05950 rickettsia
51	46	26.9	807	3	O06479 saccharomyc
52	46	26.9	1114	10	O49345 arabidopsis
53	46	26.9	1291	10	Q38800 arabidopsis
54	46	26.9	1388	5	Q18566 caenorhabdi
55	45.5	26.6	434	11	O70596 rattus norv
56	45.5	26.6	458	2	O86908 spingomona
57	45.5	26.6	534	12	O90731 feline herp
58	45	26.3	177	5	O61695 manduca sex
59	45	26.3	330	5	O16227 caenorhabdi
60	45	26.3	367	2	O51731 borrelia bu
61	45	26.3	395	2	O45381 bordetella
62	45	26.3	395	2	O87998 bordetella
63	45	26.3	396	2	O52552 prochloroco
64	45	26.3	398	2	O9WZP7 thermotoga
65	45	26.3	713	12	Q9YR00 ranid herpe
66	45	26.3	808	5	Q9XXU1 caenorhabdi
67	45	26.3	928	13	Q98906 gallus gall
68	45	26.3	1194	4	O94965 homo sapien
69	45	26.3	3429	5	Q24593 drosophila
70	44.5	26.0	400	5	O76147 dugesia jap
71	44.5	26.0	443	3	O59940 neurospora
72	44.5	26.0	581	2	O51662 borrelia bu
73	44.5	26.0	1426	2	Q9X3P6 caldicellul
74	44	25.7	191	6	Q9XSM0 ovis aries
75	44	25.7	246	11	O61435 mus musculus
76	44	25.7	327	5	O62500 caenorhabdi
77	44	25.7	339	10	Q9XIN9 arabidopsis
78	44	25.7	354	13	O57684 xenopus lae
79	44	25.7	385	13	O57676 xenopus lae
80	44	25.7	394	13	O57681 xenopus lae
81	44	25.7	427	13	O57680 xenopus lae
82	44	25.7	453	2	O51881 buchnera ap
83	44	25.7	491	1	O29081 archaeglob
84	44	25.7	723	5	O97353 toxoplasma
85	44	25.7	1023	12	O92787 bovine aden
86	44	25.7	2210	12	O66802 ebola virus
87	44	25.7	2212	12	O39794 ebola virus
88	43.5	25.4	231	2	P96138 thermus aqu
89	43.5	25.4	398	2	O31665 bacillus su
90	43.5	25.4	829	12	O39233 rhesus monk
91	43.5	25.4	829	12	Q9WRT9 macaca mula
92	43.5	25.4	890	3	O60040 ajellomyces

93	43.5	25.4	977	4	O94933	O94933 homo sapien	166	41.5	24.3	272	13	O13023	O13023 xenopus lae
94	43	25.1	78	2	P95455	P95455 pseudomonas	167	41.5	24.3	285	4	O95865	O95865 homo sapien
95	43	25.1	134	1	O9VED6	O9VED6 aeropyrum p	168	41.5	24.3	290	10	O04133	O04133 glycine max
96	43	25.1	166	1	O9YCD6	O9YCD6 aeropyrum p	169	41.5	24.3	319	12	O81263	O81263 hepatitis c
97	43	25.1	210	6	O28273	O28273 canis famill	170	41.5	24.3	369	5	O62483	O62483 caenorhabdi
98	43	25.1	288	1	O58591	O58591 methanococ	171	41.5	24.3	401	1	O59044	O59044 pyrococcus
99	43	25.1	320	2	O44215	O44215 anabaena sp	172	41.5	24.3	419	4	O9Y6K9	O9Y6K9 homo sapien
100	43	25.1	322	10	O9ZT33	O9ZT33 arabidopsis	173	41.5	24.3	430	1	O53132	O53132 methanosarc
101	43	25.1	327	5	O17130	O17130 caenorhabdi	174	41.5	24.3	517	2	O9ZD68	O9ZD68 rickettsia
102	43	25.1	334	10	O40686	O40686 oryza sativ	175	41.5	24.3	654	5	O01436	O01436 caenorhabdi
103	43	25.1	362	5	O62142	O62142 caenorhabdi	176	41.5	24.3	808	12	O86281	O86281 rice ragged
104	43	25.1	383	1	O50225	O50225 methanosarc	177	41.5	24.3	925	5	O20099	O20099 caenorhabdi
105	43	25.1	383	1	O50248	O50248 methanosarc	178	41.5	24.3	1332	5	O45599	O45599 caenorhabdi
106	43	25.1	391	2	O52822	O52822 amycolatops	179	41.5	24.3	1839	5	O17383	O17383 caenorhabdi
107	43	25.1	441	8	O33262	O33262 veronica ca	180	41	24.0	67	12	O98477	O98477 paramecium
108	43	25.1	492	12	O65571	O65571 bovine herp	181	41	24.0	164	5	O20301	O20301 caenorhabdi
109	43	25.1	533	4	O95409	O95409 homo sapien	182	41	24.0	166	2	O9WYL4	O9WYL4 thermotoga
110	43	25.1	582	10	O9XIW5	O9XIW5 oryza sativ	183	41	24.0	176	1	O58049	O58049 pyrococcus
111	43	25.1	602	5	O44787	O44787 caenorhabdi	184	41	24.0	191	12	O68130	O68130 hepatitis c
112	43	25.1	606	13	O9YGW5	O9YGW5 gallus gall	185	41	24.0	194	4	O9Y4T8	O9Y4T8 homo sapien
113	43	25.1	651	5	O44906	O44906 caenorhabdi	186	41	24.0	205	12	O68769	O68769 hepatitis c
114	43	25.1	719	11	O61281	O61281 mus musculus	187	41	24.0	241	1	O58809	O58809 methanococ
115	43	25.1	1128	11	O88442	O88442 mus musculus	188	41	24.0	251	12	P89128	P89128 sida golden
116	43	25.1	1435	5	O44937	O44937 caenorhabdi	189	41	24.0	311	10	O65189	O65189 hordeum vul
117	43	25.1	1879	5	O18210	O18210 caenorhabdi	190	41	24.0	312	2	O68338	O68338 vibrio chol
118	43	25.1	2254	10	O38971	O38971 arabidopsis	191	41	24.0	329	10	O9ZT35	O9ZT35 arabidopsis
119	43	25.1	2254	10	O38970	O38970 arabidopsis	192	41	24.0	332	10	P93519	P93519 zea mays (m
120	42.5	24.9	114	1	O9YDX0	O9YDX0 aeropyrum p	193	41	24.0	334	10	O64938	O64938 hordeum vul
121	42.5	24.9	129	2	P72383	P72383 sulfobacill	194	41	24.0	336	5	O9XEN5	O9XEN5 triticum ae
122	42.5	24.9	177	12	O41985	O41985 maize rayad	195	41	24.0	335	5	O44908	O44908 caenorhabdi
123	42.5	24.9	177	12	O41986	O41986 maize rayad	196	41	24.0	339	5	O18069	O18069 caenorhabdi
124	42.5	24.9	332	10	O9ZNY9	O9ZNY9 oryza sativ	197	41	24.0	350	5	O45803	O45803 caenorhabdi
125	42.5	24.9	462	10	O9XE94	O9XE94 glycine max	198	41	24.0	372	10	O9XFW9	O9XFW9 cicer ariet
126	42.5	24.9	466	10	O81335	O81335 mesembryant	199	41	24.0	377	3	O74794	O74794 schizosacch
127	42.5	24.9	555	3	O03359	O03359 saccharomyc	200	41	24.0	380	2	O67225	O67225 aquifex aeo
128	42.5	24.9	766	2	O33466	O33466 pseudomonas	201	41	24.0	389	13	O9YGU5	O9YGU5 brachydanio
129	42.5	24.9	766	2	O9WW20	O9WW20 pseudomonas	202	41	24.0	403	3	O42963	O42963 schizosacch
130	42.5	24.9	773	4	O15052	O15052 homo sapien	203	41	24.0	404	13	O9YGU6	O9YGU6 brachydanio
131	42.5	24.9	857	12	P90249	P90249 human immun	204	41	24.0	410	13	O9YGU8	O9YGU8 brachydanio
132	42.5	24.9	1802	5	O17163	O17163 brugia mala	205	41	24.0	434	3	P78918	P78918 schizosacch
133	42	24.6	108	12	O66011	O66011 cocksfoot m	206	41	24.0	441	8	O33204	O33204 tetrachondr
134	42	24.6	120	6	O46642	O46642 equus burch	207	41	24.0	455	10	O39248	O39248 arabidopsis
135	42	24.6	169	5	O9XYT6	O9XYT6 cassiopea x	208	41	24.0	465	10	P92979	P92979 arabidopsis
136	42	24.6	191	6	O9XS55	O9XS55 canis famill	209	41	24.0	465	10	O48886	O48886 arabidopsis
137	42	24.6	229	5	O62545	O62545 hydra atten	210	41	24.0	473	2	O85040	O85040 thioabacillu
138	42	24.6	269	13	O98873	O98873 brachydanio	211	41	24.0	483	11	O62819	O62819 rattus norv
139	42	24.6	276	12	O89621	O89621 avian adeno	212	41	24.0	507	13	O9W7K3	O9W7K3 brachydanio
140	42	24.6	289	1	O9YBP2	O9YBP2 aeropyrum p	213	41	24.0	554	11	O62390	O62390 mus musculu
141	42	24.6	320	2	O9ZAK3	O9ZAK3 anabaena va	214	41	24.0	568	5	O15968	O15968 drosophila
142	42	24.6	326	5	P91275	P91275 caenorhabdi	215	41	24.0	600	2	O9XC29	O9XC29 acinetobact
143	42	24.6	335	10	O82716	O82716 triticum ae	216	41	24.0	645	2	O54315	O54315 thermus sp.
144	42	24.6	357	5	O16397	O16397 caenorhabdi	217	41	24.0	660	2	O66925	O66925 aquifex aeo
145	42	24.6	441	8	O37180	O37180 callitriche	218	41	24.0	673	5	O76649	O76649 caenorhabdi
146	42	24.6	473	2	O9X520	O9X520 thioabacillu	219	41	24.0	704	5	O61106	O61106 trypanosoma
147	42	24.6	475	5	O46038	O46038 drosophila	220	41	24.0	799	10	O23190	O23190 arabidopsis
148	42	24.6	509	8	O9XMR6	O9XMR6 tetralymena	221	41	24.0	811	2	O66857	O66857 bradyrhizob
149	42	24.6	531	5	O21603	O21603 caenorhabdi	222	41	24.0	934	2	O84303	O84303 chlamydia t
150	42	24.6	574	3	O12555	O12555 emericella	223	41	24.0	1073	4	O94881	O94881 homo sapien
151	42	24.6	581	10	O9ZWJ1	O9ZWJ1 ipomoea bat	224	41	24.0	1095	4	O99458	O99458 homo sapien
152	42	24.6	650	5	O02306	O02306 caenorhabdi	225	41	24.0	1135	12	O9WI73	O9WI73 hantavirus
153	42	24.6	680	12	O87024	O87024 saccharomyc	226	41	24.0	1364	11	O35442	O35442 mus musculu
154	42	24.6	722	2	O92764	O92764 chlamydia p	227	41	24.0	1999	4	O99940	O99940 homo sapien
155	42	24.6	767	2	O25395	O25395 helicobacte	228	41	24.0	2003	4	O00306	O00306 homo sapien
156	42	24.6	893	5	O9Y1Y3	O9Y1Y3 ephydatia f	229	41	24.0	2523	2	O53393	O53393 mycobacteri
157	42	24.6	1505	12	O87022	O87022 saccharomyc	230	41	24.0	3014	12	O39928	O39928 hepatitis c
158	42	24.6	1509	11	O61194	O61194 mus musculus	231	40.5	23.7	139	2	O50204	O50204 xanthobacte
159	42	24.6	1658	11	O61182	O61182 mus musculus	232	40.5	23.7	144	2	O55181	O55181 synecocyst
160	42	24.6	1686	4	O00443	O00443 homo sapien	233	40.5	23.7	171	12	O83588	O83588 measles vir
161	42	24.6	2408	4	O92566	O92566 homo sapien	234	40.5	23.7	229	2	O55050	O55050 shigella so
162	42	24.6	2733	12	O39226	O39226 murine hepa	235	40.5	23.7	296	12	O9YLM7	O9YLM7 coxsackievi
163	41.5	24.3	144	1	O28617	O28617 archaeoglob	236	40.5	23.7	315	10	O9XGC2	O9XGC2 vitis vinif
164	41.5	24.3	193	12	O81216	O81216 hepatitis c	237	40.5	23.7	367	1	O58609	O58609 pyrococcus
165	41.5	24.3	193	12	O81468	O81468 hepatitis c	238	40.5	23.7	407	12	O98384	O98384 measles vir

239	40.5	23.7	473	5	Q9Y0S9	Q9Y0S9 drosophila	312	40	23.4	750	4	043748	043748 homo sapien
240	40.5	23.7	542	4	043507	043507 homo sapien	313	40	23.4	788	4	Q9Y4X7	Q9Y4X7 homo sapien
241	40.5	23.7	554	4	015245	015245 homo sapien	314	40	23.4	827	5	Q21055	Q21055 caenorhabdi
242	40.5	23.7	554	4	015395	015395 homo sapien	315	40	23.4	998	2	Q47262	Q47262 escherichia
243	40.5	23.7	555	4	Q9Y6F0	Q9Y6F0 homo sapien	316	40	23.4	998	2	P71223	P71223 escherichia
244	40.5	23.7	556	4	Q45232	Q45232 bradyrhizob	317	40	23.4	998	2	Q47461	Q47461 escherichia
245	40.5	23.7	572	3	Q12650	Q12650 penicillium	318	40	23.4	998	2	Q85101	Q85101 escherichia
246	40.5	23.7	617	12	Q90381	Q90381 measles vir	319	40	23.4	998	2	Q46716	Q46716 escherichia
247	40.5	23.7	1386	4	Q75064	Q75064 homo sapien	320	40	23.4	1012	10	Q9ZQJ7	Q9ZQJ7 arabidopsis
248	40	23.4	130	2	P70743	P70743 acidimicrob	321	40	23.4	1055	4	Q53348	Q53348 mycobacteri
249	40	23.4	143	10	Q41540	Q41540 triticum ae	322	40	23.4	1120	2	O95855	O95855 homo sapien
250	40	23.4	148	12	Q9YVW6	Q9YVW6 melanoplus	323	40	23.4	1355	4	Q75049	Q75049 homo sapien
251	40	23.4	158	5	Q21265	Q21265 caenorhabdi	324	40	23.4	1463	11	O55124	O55124 mus musculu
252	40	23.4	162	1	Q9YCX4	Q9YCX4 aetopyrum p	325	40	23.4	1795	5	Q76894	Q76894 drosophila
253	40	23.4	162	5	P91606	P91606 drosophila	326	40	23.4	1910	5	Q22426	Q22426 caenorhabdi
254	40	23.4	173	1	Q00396	Q00396 methanococ	327	40	23.4	3066	11	Q62388	Q62388 mus musculu
255	40	23.4	173	5	P91603	P91603 drosophila	328	40	23.4	4574	4	Q16640	Q16640 homo sapien
256	40	23.4	176	11	Q9WU23	Q9WU23 mus musculu	329	40	23.4	4684	4	Q15149	Q15149 homo sapien
257	40	23.4	215	1	Q28955	Q28955 archaeoglob	330	39.5	23.1	89	10	Q9XFT9	Q9XFT9 chlamydomon
258	40	23.4	220	2	O55813	O55813 synecocyst	331	39.5	23.1	131	12	Q9YF99	Q9YF99 retroperito
259	40	23.4	240	1	Q27996	Q27996 archaeoglob	332	39.5	23.1	171	12	Q83595	Q83595 measles vir
260	40	23.4	246	10	Q49011	Q49011 glycine max	333	39.5	23.1	184	5	O02048	O02048 caenorhabdi
261	40	23.4	246	10	Q49015	Q49015 glycine max	334	39.5	23.1	193	12	Q81215	Q81215 hepatitis c
262	40	23.4	251	12	P89123	P89123 sida golden	335	39.5	23.1	207	12	Q86637	Q86637 eastern equ
263	40	23.4	255	4	O95205	O95205 homo sapien	336	39.5	23.1	265	10	Q9XEL7	Q9XEL7 oryza sativ
264	40	23.4	262	2	P74577	P74577 synecocyst	337	39.5	23.1	273	5	Q93236	Q93236 caenorhabdi
265	40	23.4	267	12	Q9YR99	Q9YR99 avian adeno	338	39.5	23.1	316	5	O16463	O16463 caenorhabdi
266	40	23.4	276	1	O59350	O59350 pyrococcus	339	39.5	23.1	317	5	O16235	O16235 caenorhabdi
267	40	23.4	289	1	O29463	O29463 archaeoglob	340	39.5	23.1	326	5	Q9XYF7	Q9XYF7 drosophila
268	40	23.4	324	4	Q9Y310	Q9Y310 homo sapien	341	39.5	23.1	327	5	O76523	O76523 drosophila
269	40	23.4	329	2	O92826	O92826 chlamydia p	342	39.5	23.1	345	5	O22461	O22461 caenorhabdi
270	40	23.4	332	2	O52371	O52371 rhizobium t	343	39.5	23.1	378	1	Q33910	Q33910 pyrobaculum
271	40	23.4	332	5	O45403	O45403 caenorhabdi	344	39.5	23.1	400	1	O30076	O30076 archaeoglob
272	40	23.4	334	10	Q9XEN7	Q9XEN7 triticum ae	345	39.5	23.1	407	12	Q98391	Q98391 measles vir
273	40	23.4	336	10	Q9ZN22	Q9ZN22 oryza sativ	346	39.5	23.1	407	12	Q98377	Q98377 measles vir
274	40	23.4	359	4	Q13394	Q13394 homo sapien	347	39.5	23.1	407	12	Q98380	Q98380 measles vir
275	40	23.4	359	11	O70299	O70299 mus musculu	348	39.5	23.1	407	12	Q98381	Q98381 measles vir
276	40	23.4	361	13	O42442	O42442 oryzias lat	349	39.5	23.1	407	12	Q98383	Q98383 measles vir
277	40	23.4	373	1	O93651	O93651 archaeoglob	350	39.5	23.1	407	12	Q98386	Q98386 measles vir
278	40	23.4	386	13	Q93370	Q93370 brachydanio	351	39.5	23.1	407	12	Q98389	Q98389 measles vir
279	40	23.4	388	2	Q9XDW8	Q9XDW8 streptococ	352	39.5	23.1	407	12	Q98390	Q98390 measles vir
280	40	23.4	422	10	Q38948	Q38948 arabidopsis	353	39.5	23.1	407	12	Q98392	Q98392 measles vir
281	40	23.4	437	5	Q19304	Q19304 caenorhabdi	354	39.5	23.1	407	12	Q98393	Q98393 measles vir
282	40	23.4	441	8	Q32100	Q32100 callitriche	355	39.5	23.1	407	12	Q98378	Q98378 measles vir
283	40	23.4	445	12	Q9YKJ7	Q9YKJ7 hepatitis b	356	39.5	23.1	407	12	Q98379	Q98379 measles vir
284	40	23.4	448	8	Q32402	Q32402 hyparrhenia	357	39.5	23.1	407	12	Q98382	Q98382 measles vir
285	40	23.4	448	8	Q33185	Q33185 tristachya	358	39.5	23.1	407	12	Q98385	Q98385 measles vir
286	40	23.4	453	5	O96420	O96420 drosophila	359	39.5	23.1	407	12	Q98387	Q98387 measles vir
287	40	23.4	458	10	O48887	O48887 arabidopsis	360	39.5	23.1	407	12	Q98388	Q98388 measles vir
288	40	23.4	458	10	P92980	P92980 arabidopsis	361	39.5	23.1	434	10	Q9XG11	Q9XG11 lycopersico
289	40	23.4	459	5	P91876	P91876 drosophila	362	39.5	23.1	445	5	O17204	O17204 caenorhabdi
290	40	23.4	464	10	Q92P22	Q92P22 brassica ju	363	39.5	23.1	485	5	Q9XW17	Q9XW17 caenorhabdi
291	40	23.4	465	8	O98704	O98704 polydremum	364	39.5	23.1	486	2	Q92134	Q92134 bradyrhizob
292	40	23.4	465	8	O98943	O98943 gelidella	365	39.5	23.1	495	11	O35795	O35795 rattus norv
293	40	23.4	473	2	Q92H21	Q92H21 thobacillu	366	39.5	23.1	579	5	O61264	O61264 caenorhabdi
294	40	23.4	473	2	Q9XD77	Q9XD77 nitrobacter	367	39.5	23.1	581	13	Q93274	Q93274 xenopus lae
295	40	23.4	473	2	Q9XD76	Q9XD76 nitrobacter	368	39.5	23.1	598	5	O16901	O16901 caenorhabdi
296	40	23.4	473	8	O78385	O78385 verbena off	369	39.5	23.1	613	12	Q83583	Q83583 measles vir
297	40	23.4	485	5	Q9XTV6	Q9XTV6 caenorhabdi	370	39.5	23.1	617	5	Q19844	Q19844 caenorhabdi
298	40	23.4	494	2	O33767	O33767 sphingomona	371	39.5	23.1	617	12	Q83639	Q83639 measles vir
299	40	23.4	505	2	O85980	O85980 sphingomona	372	39.5	23.1	617	12	O42032	O42032 measles vir
300	40	23.4	512	4	O60678	O60678 homo sapien	373	39.5	23.1	617	12	Q83636	Q83636 measles vir
301	40	23.4	528	11	O70467	O70467 rattus norv	374	39.5	23.1	617	12	Q83638	Q83638 measles vir
302	40	23.4	589	2	O06827	O06827 mycobacteri	375	39.5	23.1	617	12	Q83641	Q83641 measles vir
303	40	23.4	632	13	Q90356	Q90356 cyprinus ca	376	39.5	23.1	617	12	Q83642	Q83642 measles vir
304	40	23.4	663	5	O61910	O61910 caenorhabdi	377	39.5	23.1	617	12	Q83644	Q83644 measles vir
305	40	23.4	666	2	O33681	O33681 streptococ	378	39.5	23.1	617	12	O41002	O41002 measles vir
306	40	23.4	666	2	Q54780	Q54780 streptococ	379	39.5	23.1	617	12	O41003	O41003 measles vir
307	40	23.4	683	8	Q9XPS8	Q9XPS8 triticum ae	380	39.5	23.1	617	12	O41004	O41004 measles vir
308	40	23.4	724	5	O17856	O17856 caenorhabdi	381	39.5	23.1	617	12	Q83297	Q83297 measles vir
309	40	23.4	728	6	O97567	O97567 bos taurus	382	39.5	23.1	617	12	Q03908	Q03908 measles vir
310	40	23.4	731	4	O43517	O43517 homo sapien	383	39.5	23.1	617	12	Q03911	Q03911 measles vir
311	40	23.4	732	11	P70605	P70605 rattus norv	384	39.5	23.1	617	12	Q03912	Q03912 measles vir

385	39.5	23.1	617	12	Q83635	measles vir	458	39	22.8	174	5	Q96624	dictyosteli
386	39.5	23.1	617	12	Q83632	measles vir	459	39	22.8	195	5	Q18278	caenorhabdi
387	39.5	23.1	617	12	P90460	measles vir	460	39	22.8	198	12	Q65591	bovine immu
388	39.5	23.1	617	12	Q9WJ00	measles vir	461	39	22.8	198	12	Q65594	bovine immu
389	39.5	23.1	617	12	Q9WJ29	measles vir	462	39	22.8	215	10	Q41061	pisum sativ
390	39.5	23.1	617	12	Q9WJ28	measles vir	463	39	22.8	215	5	Q76586	caenorhabdi
391	39.5	23.1	617	12	Q9WJ27	measles vir	464	39	22.8	247	8	Q99378	sepietta ow
392	39.5	23.1	617	12	Q9WJ26	measles vir	465	39	22.8	252	13	Q57679	xenopus lae
393	39.5	23.1	617	12	Q9WJ25	measles vir	466	39	22.8	255	2	Q53084	rhodococcus
394	39.5	23.1	617	12	Q9WJ24	measles vir	467	39	22.8	266	2	Q54509	erwinia any
395	39.5	23.1	617	12	Q9WJ23	measles vir	468	39	22.8	281	13	Q73821	caenorhabdi
396	39.5	23.1	617	12	Q9WJ22	measles vir	469	39	22.8	285	2	Q9X5X8	pseudomonas
397	39.5	23.1	617	12	Q9WJ21	measles vir	470	39	22.8	288	2	Q07487	versinia en
398	39.5	23.1	617	12	Q9WJ20	measles vir	471	39	22.8	290	13	Q57678	caenorhabdi
399	39.5	23.1	617	12	Q9WJ19	measles vir	472	39	22.8	292	13	Q9Y942	brachydanio
400	39.5	23.1	617	12	Q9WJ18	measles vir	473	39	22.8	297	8	Q03301	chelus flmb
401	39.5	23.1	617	12	Q9WJ17	measles vir	474	39	22.8	299	2	Q9X259	thermotoga
402	39.5	23.1	617	12	Q9WJ16	measles vir	475	39	22.8	301	2	Q9X8P7	streptomyce
403	39.5	23.1	617	12	Q9WJ15	measles vir	476	39	22.8	307	5	Q21109	caenorhabdi
404	39.5	23.1	617	12	Q9WJ14	measles vir	477	39	22.8	310	4	Q13460	homo sapien
405	39.5	23.1	617	12	Q9WJ13	measles vir	478	39	22.8	325	1	Q9Y942	aeropyrum p
406	39.5	23.1	617	12	Q9WJ12	measles vir	479	39	22.8	328	3	Q07379	saccharomyc
407	39.5	23.1	617	12	Q9WJ11	measles vir	480	39	22.8	329	5	Q18071	caenorhabdi
408	39.5	23.1	617	12	Q9WJ10	measles vir	481	39	22.8	329	13	Q57677	caenorhabdi
409	39.5	23.1	617	12	Q9WJ09	measles vir	482	39	22.8	331	10	Q92P12	cicer ariet
410	39.5	23.1	617	12	Q9WJ08	measles vir	483	39	22.8	332	5	Q22982	caenorhabdi
411	39.5	23.1	617	12	Q9WJ07	measles vir	484	39	22.8	339	5	Q17996	caenorhabdi
412	39.5	23.1	617	12	Q9WJ06	measles vir	485	39	22.8	342	5	Q21978	caenorhabdi
413	39.5	23.1	617	12	Q9WJ05	measles vir	486	39	22.8	351	13	Q9Y909	brachydanio
414	39.5	23.1	617	12	Q9WJ04	measles vir	487	39	22.8	357	11	Q63785	rattus norv
415	39.5	23.1	617	12	Q9WJ03	measles vir	488	39	22.8	364	5	Q19494	caenorhabdi
416	39.5	23.1	617	12	Q9WJ02	measles vir	489	39	22.8	364	5	Q44442	caenorhabdi
417	39.5	23.1	617	12	Q9WJ01	measles vir	490	39	22.8	370	4	Q43311	homo sapien
418	39.5	23.1	617	12	Q9W007	measles vir	491	39	22.8	376	2	Q9XC59	pseudomonas
419	39.5	23.1	617	12	Q9W9G6	measles vir	492	39	22.8	380	11	Q63784	rattus norv
420	39.5	23.1	617	12	Q9W8V2	measles vir	493	39	22.8	386	13	Q90929	gallus gall
421	39.5	23.1	617	12	Q9W8G7	measles vir	494	39	22.8	388	4	Q43797	homo sapien
422	39.5	23.1	617	12	Q9W8B0	measles vir	495	39	22.8	398	3	Q13343	rhodotorula
423	39.5	23.1	617	12	Q9W882	measles vir	496	39	22.8	407	12	Q98376	measles vir
424	39.5	23.1	625	5	Q94263	caenorhabdi	497	39	22.8	425	2	Q50347	lactobacill
425	39.5	23.1	701	4	Q75342	homo sapien	498	39	22.8	438	2	Q52582	staphylococ
426	39.5	23.1	718	10	Q80771	arabidopsis	499	39	22.8	440	8	Q98382	triatolola pi
427	39.5	23.1	745	5	Q21706	caenorhabdi	500	39	22.8	441	8	Q37225	hippuris vu
428	39.5	23.1	767	2	Q9Z1F4	helicobacte	501	39	22.8	441	8	Q98388	minulus aur
429	39.5	23.1	1051	4	Q95135	homo sapien	502	39	22.8	444	11	Q62219	mus musculu
430	39.5	23.1	1187	13	Q12965	fugu rubrip	503	39	22.8	454	13	Q91675	caenorhabdi
431	39.5	23.1	1241	12	Q66579	eastern equ	504	39	22.8	455	10	Q64510	arabidopsis
432	39.5	23.1	1242	12	Q88678	eastern equ	505	39	22.8	459	5	Q61268	drosophi
433	39.5	23.1	1242	12	Q88790	eastern equ	506	39	22.8	460	13	Q91728	caenorhabdi
434	39.5	23.1	1242	12	Q88792	eastern equ	507	39	22.8	461	13	Q90928	gallus gall
435	39.5	23.1	1242	12	Q88793	eastern equ	508	39	22.8	462	13	Q57682	caenorhabdi
436	39.5	23.1	1242	12	Q88794	eastern equ	509	39	22.8	464	2	Q9WJ25	streptomyce
437	39.5	23.1	1242	12	Q88795	eastern equ	510	39	22.8	466	11	Q70251	mus musculu
438	39.5	23.1	1242	12	Q88796	eastern equ	511	39	22.8	467	1	Q27845	methanobact
439	39.5	23.1	1242	12	Q88797	eastern equ	512	39	22.8	467	8	Q20226	hebenstreti
440	39.5	23.1	1242	12	Q88798	eastern equ	513	39	22.8	467	8	Q98662	ecreocarp
441	39.5	23.1	1242	12	Q88799	eastern equ	514	39	22.8	468	8	Q98383	halleria lu
442	39.5	23.1	1242	12	Q88359	eastern equ	515	39	22.8	469	2	Q9X970	synecococc
443	39.5	23.1	1569	5	Q97234	plasmodium	516	39	22.8	469	8	Q46893	schwenckia
444	39.5	23.1	1732	5	Q09451	bonnemaiso	517	39	22.8	470	2	Q69686	mycobacteri
445	39.5	23.1	2490	4	Q12923	homo sapien	518	39	22.8	472	3	Q13648	schizosacch
446	39.5	23.1	2513	12	Q90368	o'nyong-nyo	519	39	22.8	473	2	Q92B35	pseudomonas
447	39.5	23.1	2513	12	Q90370	igbo ora vi	520	39	22.8	476	5	Q61267	leporid her
448	39.5	23.1	2685	3	Q14151	schizosacch	521	39	22.8	483	12	Q82172	caenorhabdi
449	39	22.8	77	12	Q9WT71	tt virus. o	522	39	22.8	487	11	Q63783	rattus norv
450	39	22.8	77	12	Q9WT69	tt virus. o	523	39	22.8	489	5	Q9XN3	caenorhabdi
451	39	22.8	104	2	Q33291	mycobacteri	524	39	22.8	490	10	Q22162	arabidopsis
452	39	22.8	108	2	Q9ZB54	streptomyce	525	39	22.8	493	5	Q17534	caenorhabdi
453	39	22.8	113	12	Q70647	simian t-ce	526	39	22.8	497	13	Q57685	caenorhabdi
454	39	22.8	135	12	Q84655	paramecium	527	39	22.8	498	11	Q63573	caenorhabdi
455	39	22.8	142	11	Q9Z0V9	mus musculu	528	39	22.8	498	13	Q90927	rattus norv
456	39	22.8	166	12	Q66051	herpesvirus	529	39	22.8	506	1	P95942	gallus gall
457	39	22.8	168	5	P91710	herpesvirus	530	39	22.8	507	5	Q9XVK6	sulfolobus
						P91710 drosophila							caenorhabdi

531	39	22.8	509	11	P70250	P70250 mus musculus	604	38.5	22.5	387	2	O32062	O32062 bacillus su
532	39	22.8	509	13	Q90925	Q90925 gallus gall	605	38.5	22.5	439	4	O75967	O75967 homo sapien
533	39	22.8	511	10	O64631	O64631 arabidopsis	606	38.5	22.5	443	4	O95967	O95967 homo sapien
534	39	22.8	514	5	Q21587	Q21587 caenorhabdi	607	38.5	22.5	443	11	O55058	O55058 cricetus
535	39	22.8	521	3	O08777	O08777 saccharomyc	608	38.5	22.5	471	12	Q9WNM5	Q9WNM5 human papil
536	39	22.8	530	4	O14928	O14928 homo sapien	609	38.5	22.5	527	2	Q9Z9S2	Q9Z9S2 bacillus sp
537	39	22.8	531	11	P70624	P70624 rattus norv	610	38.5	22.5	533	3	O13490	O13490 myrothecium
538	39	22.8	532	11	O02780	O02780 mus musculus	611	38.5	22.5	541	2	P71370	P71370 haemophilus
539	39	22.8	533	2	O52571	O52571 amycolatops	612	38.5	22.5	555	13	Q90874	Q90874 gallus gall
540	39	22.8	534	13	Q9Y156	Q9Y156 gallus gall	613	38.5	22.5	562	5	Q18810	Q18810 caenorhabdi
541	39	22.8	547	2	O84216	O84216 chlamydia t	614	38.5	22.5	574	4	O43734	O43734 homo sapien
542	39	22.8	574	3	Q9Y7P6	Q9Y7P6 schizosacch	615	38.5	22.5	600	12	Q9YNB6	Q9YNB6 measles vir
543	39	22.8	581	2	O34023	O34023 chlamydia p	616	38.5	22.5	607	12	O91244	O91244 measles vir
544	39	22.8	600	12	O91245	O91245 measles vir	617	38.5	22.5	617	12	O9WJ19	O9WJ19 measles vir
545	39	22.8	601	5	O77459	O77459 drosophila	618	38.5	22.5	617	12	O56667	O56667 measles vir
546	39	22.8	604	3	O13865	O13865 schizosacch	619	38.5	22.5	617	12	O56668	O56668 measles vir
547	39	22.8	612	2	O69801	O69801 streptomyce	620	38.5	22.5	617	12	O56669	O56669 measles vir
548	39	22.8	614	3	Q01203	Q01203 melampora	621	38.5	22.5	617	12	O56670	O56670 measles vir
549	39	22.8	716	12	O9YTP4	O9YTP4 ateline her	622	38.5	22.5	617	12	O56671	O56671 measles vir
550	39	22.8	732	2	O86424	O86424 enterobacte	623	38.5	22.5	617	12	O56672	O56672 measles vir
551	39	22.8	752	3	O13400	O13400 ustilago ma	624	38.5	22.5	617	12	O56673	O56673 measles vir
552	39	22.8	755	4	O75037	O75037 homo sapien	625	38.5	22.5	617	12	O56675	O56675 measles vir
553	39	22.8	757	3	O13399	O13399 ustilago ma	626	38.5	22.5	617	12	O57282	O57282 measles vir
554	39	22.8	782	11	Q9WU23	Q9WU23 mus musculus	627	38.5	22.5	631	12	O57292	O57292 measles vir
555	39	22.8	829	12	P87584	P87584 chipmunk pa	628	38.5	22.5	631	11	O70418	O70418 rattus norv
556	39	22.8	893	5	O76331	O76331 drosophila	629	38.5	22.5	638	12	O89399	O89399 paramencium
557	39	22.8	950	13	O93508	O93508 brachydanio	630	38.5	22.5	647	4	O75671	O75671 homo sapien
558	39	22.8	955	5	O45195	O45195 caenorhabdi	631	38.5	22.5	651	11	Q03719	Q03719 mus musculus
559	39	22.8	971	5	O45785	O45785 caenorhabdi	632	38.5	22.5	712	10	O80773	O80773 arabidopsis
560	39	22.8	1013	10	O04451	O04451 arabidopsis	633	38.5	22.5	788	12	P88932	P88932 kaposi's sa
561	39	22.8	1022	5	O61907	O61907 caenorhabdi	634	38.5	22.5	788	12	O40933	O40933 kaposi's sa
562	39	22.8	1041	4	Q9Y441	Q9Y441 homo sapien	635	38.5	22.5	856	2	Q9ZFG9	Q9ZFG9 azotobacter
563	39	22.8	1127	12	Q89839	Q89839 hantaan vir	636	38.5	22.5	968	5	P91289	P91289 caenorhabdi
564	39	22.8	1135	12	O67794	O67794 hantaan vir	637	38.5	22.5	1012	12	P88907	P88907 kaposi's sa
565	39	22.8	1135	12	O91736	O91736 hantaan vir	638	38.5	22.5	1012	12	O40915	O40915 kaposi's sa
566	39	22.8	1135	12	O67795	O67795 hantaan vir	639	38.5	22.5	1053	12	Q9WSQ2	Q9WSQ2 african hor
567	39	22.8	1135	12	O67796	O67796 hantaan vir	640	38.5	22.5	1110	5	O09493	O09493 caenorhabdi
568	39	22.8	1158	4	Q14113	Q14113 homo sapien	641	38.5	22.5	1237	2	O84549	O84549 chlamydia t
569	39	22.8	1224	13	Q9Y190	Q9Y190 xenopus lae	642	38.5	22.5	1275	12	O41977	O41977 murine herp
570	39	22.8	1357	5	Q22453	Q22453 caenorhabdi	643	38.5	22.5	1534	12	O88627	O88627 borna disea
571	39	22.8	1373	11	O70438	O70438 mus musculus	644	38.5	22.5	1758	5	O19098	O19098 caenorhabdi
572	39	22.8	1390	5	O17602	O17602 caenorhabdi	645	38.5	22.5	1759	5	O19099	O19099 caenorhabdi
573	39	22.8	1786	5	O17344	O17344 caenorhabdi	646	38.5	22.5	2386	3	Q9Z391	Q9Z391 schizosacch
574	39	22.8	1809	5	O17487	O17487 caenorhabdi	647	38.5	22.5	2484	6	Q28006	Q28006 bos taurus
575	39	22.8	1815	5	O17488	O17488 caenorhabdi	648	38.5	22.5	4273	2	O31781	O31781 bacillus su
576	39	22.8	1867	5	O17486	O17486 caenorhabdi	649	38.5	22.5	4545	11	Q61291	Q61291 mus musculus
577	39	22.8	1925	12	O9YRB3	O9YRB3 nudaurelia	650	38	22.2	84	12	O84466	O84466 paramencium
578	39	22.8	2039	5	O17489	O17489 caenorhabdi	651	38	22.2	86	10	O41572	O41572 triticum ae
579	39	22.8	2304	12	Q88893	Q88893 tobacco rin	652	38	22.2	92	2	Q49324	Q49324 mycoplasma
580	39	22.8	2527	5	Q24107	Q24107 drosophila	653	38	22.2	111	10	O42076	O42076 arabidopsis
581	39	22.8	2529	5	Q24605	Q24605 drosophila	654	38	22.2	129	10	O23999	O23999 hordeum vul
582	39	22.8	2864	12	O69422	O69422 hepatitis g	655	38	22.2	132	1	O9YG22	O9YG22 aeropyrum p
583	39	22.8	6994	5	O17343	O17343 caenorhabdi	656	38	22.2	138	12	O68204	O68204 hepatitis c
584	39	22.8	7576	2	Q9ZG44	Q9ZG44 streptomyce	657	38	22.2	142	11	O35558	O35558 mus musculus
585	38.5	22.5	158	12	O40910	O40910 kaposi's sa	658	38	22.2	151	2	P78186	P78186 escherichia
586	38.5	22.5	159	12	O87661	O87661 eastern equ	659	38	22.2	169	2	O07698	O07698 mycobacteri
587	38.5	22.5	192	12	O81403	O81403 hepatitis c	660	38	22.2	173	5	P91723	P91723 drosophila
588	38.5	22.5	193	12	O81214	O81214 hepatitis c	661	38	22.2	174	13	O9YHX6	O9YHX6 lampetra fl
589	38.5	22.5	196	5	O17340	O17340 caenorhabdi	662	38	22.2	177	12	O41983	O41983 maize rayad
590	38.5	22.5	235	12	O67655	O67655 garlic late	663	38	22.2	177	12	O41984	O41984 maize rayad
591	38.5	22.5	249	2	Q9ZH28	Q9ZH28 mycoplasma	664	38	22.2	177	12	O41987	O41987 maize rayad
592	38.5	22.5	267	12	O65205	O65205 african swi	665	38	22.2	177	12	O41988	O41988 maize rayad
593	38.5	22.5	291	10	O65188	O65188 zantedesch	666	38	22.2	177	12	O41990	O41990 maize rayad
594	38.5	22.5	296	12	O9YLM9	O9YLM9 coxsackievi	667	38	22.2	177	12	O41991	O41991 maize rayad
595	38.5	22.5	303	10	Q9XHY0	Q9XHY0 oryza sativ	668	38	22.2	177	12	O41992	O41992 maize rayad
596	38.5	22.5	305	5	O16869	O16869 caenorhabdi	669	38	22.2	177	12	O41994	O41994 maize rayad
597	38.5	22.5	307	2	Q52874	Q52874 rhodothermu	670	38	22.2	177	12	O41995	O41995 maize rayad
598	38.5	22.5	314	11	O62526	O62526 drosophila	671	38	22.2	177	12	O41996	O41996 maize rayad
599	38.5	22.5	328	2	O84404	O84404 chlamydia t	672	38	22.2	194	2	O50238	O50238 zymomonas m
600	38.5	22.5	348	10	Q39900	Q39900 glycine max	673	38	22.2	203	3	Q03201	Q03201 saccharomyc
601	38.5	22.5	347	1	O29274	O29274 archaeoglob	674	38	22.2	204	3	Q9Y7T8	Q9Y7T8 schizosacch
602	38.5	22.5	348	10	Q39900	Q39900 glycine max	675	38	22.2	213	2	O05880	O05880 mycobacteri
603	38.5	22.5	387	2	O52861	O52861 bacillus su	676	38	22.2	215	2	Q9Z6E1	Q9Z6E1 synchococc

677	38	22.2	218	2	Q926E2	Q926E2 synechococ	750	488	8	Q32332	Q32332 gelidium pu
678	38	22.2	221	11	P97516	P97516 phodopus su	751	488	8	Q33482	Q33482 gelidium la
679	38	22.2	228	3	Q94503	Q94503 schizosach	752	489	8	Q32329	Q32329 gelidium pu
680	38	22.2	235	13	Q42333	Q42333 cynops pyrr	753	492	10	Q8XHH3	Q8XHH3 lycopersico
681	38	22.2	244	11	O08972	O08972 mus musculu	754	501	11	Q920M2	Q920M2 cavia porce
682	38	22.2	253	2	Q48905	Q48905 microcystis	755	530	4	O60656	O60656 homo sapien
683	38	22.2	256	13	Q9YHY0	Q9YHY0 lampetra fl	756	530	4	O00473	O00473 homo sapien
684	38	22.2	259	1	O33163	O33163 methanosarc	757	530	4	O00474	O00474 homo sapien
685	38	22.2	259	1	O33167	O33167 methanosarc	758	530	6	O18777	O18777 oryctolagus
686	38	22.2	261	12	Q9YXP9	Q9YXP9 maize strea	759	531	10	Q9XFM4	Q9XFM4 agropyrum e
687	38	22.2	265	5	Q24404	Q24404 drosophila	760	541	10	O65449	O65449 arabidopsis
688	38	22.2	273	5	P91504	P91504 caenorhabdi	761	551	5	O44607	O44607 caenorhabdi
689	38	22.2	275	11	O63841	O63841 mus musculu	762	555	2	Q32382	Q32382 streptomyce
690	38	22.2	285	2	Q923Y0	Q923Y0 pseudomonas	763	555	10	O23408	O23408 arabidopsis
691	38	22.2	285	4	O13406	O13406 homo sapien	764	556	3	O42713	O42713 agaricus bi
692	38	22.2	291	10	O06446	O06446 solanum tub	765	560	3	O59954	O59954 emericea
693	38	22.2	293	10	Q9XH08	Q9XH08 arabidopsis	766	562	2	O34894	P34894 bacillus su
694	38	22.2	293	10	Q9XEX4	Q9XEX4 arabidopsis	767	563	10	P93542	P93542 sambucus ni
695	38	22.2	301	10	O65896	O65896 arabidopsis	768	563	10	O04367	O04367 sambucus ni
696	38	22.2	304	10	Q9XEX5	Q9XEX5 arabidopsis	769	565	5	Q20775	Q20775 caenorhabdi
697	38	22.2	308	5	O02296	O02296 caenorhabdi	770	565	12	O57295	O57295 sendai viru
698	38	22.2	309	1	Q9YCS8	Q9YCS8 aeropyrum p	771	565	12	O88252	O88252 sendai viru
699	38	22.2	312	5	O17242	O17242 caenorhabdi	772	565	12	O88253	O88253 sendai viru
700	38	22.2	315	3	Q9Y7V3	Q9Y7V3 schizosacch	773	565	12	O88254	O88254 sendai viru
701	38	22.2	316	5	O17931	O17931 caenorhabdi	774	565	12	O88255	O88255 sendai viru
702	38	22.2	317	4	O13407	O13407 homo sapien	775	565	12	O88256	O88256 sendai viru
703	38	22.2	320	5	Q9XV64	Q9XV64 caenorhabdi	776	565	12	O88257	O88257 sendai viru
704	38	22.2	321	5	O45795	O45795 caenorhabdi	777	565	12	O88258	O88258 sendai viru
705	38	22.2	323	2	O34009	O34009 rhodobacter	778	565	12	O88259	O88259 sendai viru
706	38	22.2	326	13	Q93396	Q93396 xenopus lae	779	565	12	O88260	O88260 sendai viru
707	38	22.2	328	5	Q9XAP3	Q9XAP3 caenorhabdi	780	565	12	O88412	O88412 sendai viru
708	38	22.2	334	5	O24403	O24403 drosophila	781	565	12	O88247	O88247 human parai
709	38	22.2	334	5	Q18640	Q18640 caenorhabdi	782	565	12	O88248	O88248 human parai
710	38	22.2	334	10	O22619	O22619 oryza sativ	783	565	12	O88267	O88267 human parai
711	38	22.2	336	5	O77121	O77121 janiodes sp	784	565	12	O88251	O88251 sendai viru
712	38	22.2	336	5	O46344	O46344 caenorhabdi	785	565	12	O9Y279	O9Y279 sendai viru
713	38	22.2	337	10	O65571	O65571 arabidopsis	786	565	12	O9YNH4	O9YNH4 sendai viru
714	38	22.2	342	2	O67713	O67713 aquifex aeo	787	565	12	O9YNH3	O9YNH3 sendai viru
715	38	22.2	345	1	O28335	O28335 archaeoglob	788	565	12	O9YNH2	O9YNH2 sendai viru
716	38	22.2	347	2	O51533	O51533 borrelia bu	789	565	12	O9YNH1	O9YNH1 sendai viru
717	38	22.2	349	5	O02333	O02333 caenorhabdi	790	565	12	O9YJP8	O9YJP8 sendai viru
718	38	22.2	362	2	O83854	O83854 treponema p	791	565	12	O9YIY9	O9YIY9 sendai viru
719	38	22.2	362	2	O9XWR5	O9XWR5 thermotoga	792	569	2	O9X7L2	O9X7L2 rhizobium m
720	38	22.2	362	5	O01493	O01493 caenorhabdi	793	576	2	O9X9Y4	O9X9Y4 streptomyce
721	38	22.2	366	1	O38946	O38946 pyrococcus	794	585	2	O66671	O66671 aquifex aeo
722	38	22.2	366	10	O49681	O49681 arabidopsis	795	586	12	O9W8B7	O9W8B7 dioscorea a
723	38	22.2	367	2	Q9X0F5	Q9X0F5 thermotoga	796	587	5	O44183	O44183 caenorhabdi
724	38	22.2	374	2	P72620	P72620 synechocyst	797	598	5	O18685	O18685 caenorhabdi
725	38	22.2	380	10	Q9XE20	Q9XE20 nicotiana t	798	603	5	O9XYA3	O9XYA3 sepi offic
726	38	22.2	388	13	Q9YH95	Q9YH95 xenopus lae	799	612	5	O20596	O20596 caenorhabdi
727	38	22.2	394	2	Q45632	Q45632 bacillus st	800	615	10	O9XIG1	O9XIG1 arabidopsis
728	38	22.2	397	2	P94400	P94400 bacillus su	801	616	10	O92Q03	O92Q03 arabidopsis
729	38	22.2	408	5	O96977	O96977 kentrophoro	802	632	2	O83428	O83428 treponema p
730	38	22.2	437	2	P96395	P96395 mycobacteri	803	633	10	O22854	O22854 arabidopsis
731	38	22.2	437	4	O00534	O00534 homo sapien	804	653	5	O97245	O97245 plasmodium
732	38	22.2	445	8	Q36629	Q36629 phyllophora	805	668	2	P72756	P72756 synechocyst
733	38	22.2	446	13	Q9YIE2	Q9YIE2 anguilla ja	806	706	5	O02166	O02166 caenorhabdi
734	38	22.2	448	10	O81600	O81600 arabidopsis	807	710	2	O83083	O83083 treponema p
735	38	22.2	451	2	Q9WY86	Q9WY86 thermotoga	808	727	4	O9Y5S6	O9Y5S6 homo sapien
736	38	22.2	454	5	Q20677	Q20677 caenorhabdi	809	745	12	Q9WSV7	Q9WSV7 tt virus. d
737	38	22.2	456	11	P97535	P97535 rattus norv	810	747	5	O76750	O76750 haemochus
738	38	22.2	461	13	Q9YIC0	Q9YIC0 oryzias lat	811	776	10	O65400	O65400 arabidopsis
739	38	22.2	462	8	O46983	O46983 vaucheria b	812	786	4	O15452	O15452 homo sapien
740	38	22.2	470	2	Q9X5S1	Q9X5S1 streptomyce	813	786	4	O15399	O15399 homo sapien
741	38	22.2	470	8	Q32316	Q32316 gelidium pu	814	815	4	O9Y5S7	O9Y5S7 homo sapien
742	38	22.2	471	8	Q33457	Q33457 gelidium at	815	843	11	P97582	P97582 rattus norv
743	38	22.2	474	8	O46982	O46982 tribonema i	816	944	3	P87114	P87114 schizosacch
744	38	22.2	474	8	O47285	O47285 bumilleriop	817	949	10	O41524	O41524 triticum ae
745	38	22.2	476	2	O67845	O67845 aquifex aeo	818	1032	5	O62380	O62380 caenorhabdi
746	38	22.2	476	8	O98616	O98616 xanthonema	819	1035	2	O21079	O21079 caenorhabdi
747	38	22.2	480	11	O88533	O88533 mus musculu	820	1062	2	O55528	O55528 synechocyst
748	38	22.2	488	8	Q32330	Q32330 gelidium pu	821	1069	12	O69140	O69140 human herpe
749	38	22.2	488	8	Q32331	Q32331 gelidium pu	822	1184	5	O77362	O77362 plasmodium

823	38	22.2	1200	5	076536	076536 strongyloce	896	37	21.6	137	3	P87275	P87275 saccharomyc
824	38	22.2	1221	5	021344	021344 caenorhabdi	897	37	21.6	145	1	Q9Y9M9	Q9Y9M9 aeropyrum p
825	38	22.2	1233	5	096197	096197 plasmodium	898	37	21.6	148	1	Q57936	Q57936 methanococc
826	38	22.2	1395	5	044924	044924 drosophila	899	37	21.6	151	5	Q9XVP0	Q9XVP0 caenorhabdi
827	38	22.2	1646	5	019338	019338 caenorhabdi	900	37	21.6	154	4	Q9Y351	Q9Y351 homo sapien
828	38	22.2	1703	11	Q92019	Q92019 mus musculu	901	37	21.6	156	10	O24552	O24552 vitis vinif
829	38	22.2	1973	5	Q17465	Q17465 caenorhabdi	902	37	21.6	164	2	Q56431	Q56431 thermus aqu
830	38	22.2	2305	5	P90749	P90749 caenorhabdi	903	37	21.6	166	12	Q68431	Q68431 heparitis c
831	38	22.2	2470	13	Q90681	Q90681 gallus gall	904	37	21.6	174	2	O34217	O34217 uncultured
832	38	22.2	2970	12	O56073	O56073 heparitis g	905	37	21.6	181	5	O17726	O17726 caenorhabdi
833	38	22.2	9376	2	O85168	O85168 pseudomonas	906	37	21.6	182	5	Q18388	Q18388 caenorhabdi
834	37.5	21.9	144	2	O52206	O52206 morganella	907	37	21.6	182	5	Q18389	Q18389 caenorhabdi
835	37.5	21.9	144	2	O52210	O52210 serratia ma	908	37	21.6	182	5	O45171	O45171 caenorhabdi
836	37.5	21.9	149	12	Q68722	Q68722 heparitis c	909	37	21.6	186	1	O59580	O59580 pyrococcus
837	37.5	21.9	169	11	Q35315	Q35315 mus musculu	910	37	21.6	188	11	Q64335	Q64335 rattus norv
838	37.5	21.9	176	1	O58119	O58119 pyrococcus	911	37	21.6	207	10	Q92WT0	Q92WT0 adiantum ca
839	37.5	21.9	193	12	Q81460	Q81460 heparitis c	912	37	21.6	221	10	O24098	O24098 medicago tr
840	37.5	21.9	193	12	Q81479	Q81479 heparitis c	913	37	21.6	230	11	O63122	O63122 rattus norv
841	37.5	21.9	247	5	Q20704	Q20704 caenorhabdi	914	37	21.6	231	5	O61479	O61479 anopheles g
842	37.5	21.9	288	2	Q929T4	Q929T4 bacillus sp	915	37	21.6	231	5	O61481	O61481 anopheles g
843	37.5	21.9	298	10	Q9XHQ7	Q9XHQ7 arabidopsis	916	37	21.6	231	5	O61487	O61487 anopheles g
844	37.5	21.9	302	11	O88725	O88725 mus musculu	917	37	21.6	231	5	O62615	O62615 anopheles g
845	37.5	21.9	304	3	O42806	O42806 aspergillus	918	37	21.6	232	5	O61482	O61482 anopheles g
846	37.5	21.9	308	5	O16503	O16503 caenorhabdi	919	37	21.6	232	5	O61489	O61489 anopheles g
847	37.5	21.9	316	13	Q90944	Q90944 gallus gall	920	37	21.6	238	10	O49013	O49013 glycine max
848	37.5	21.9	319	12	Q81478	Q81478 heparitis c	921	37	21.6	238	10	O49016	O49016 glycine max
849	37.5	21.9	332	10	Q92N21	Q92N21 oryza sativ	922	37	21.6	239	1	Q9YF02	Q9YF02 aeropyrum p
850	37.5	21.9	335	5	Q9XZP3	Q9XZP3 cupliennius	923	37	21.6	239	5	O61484	O61484 anopheles g
851	37.5	21.9	335	10	O82132	O82132 arabidopsis	924	37	21.6	239	5	O61485	O61485 anopheles g
852	37.5	21.9	336	2	Q9X0C3	Q9X0C3 thermotoga	925	37	21.6	248	1	O27872	O27872 methanobact
853	37.5	21.9	361	4	Q15595	Q15595 homo sapien	926	37	21.6	250	1	Q9Y8M7	Q9Y8M7 aeropyrum p
854	37.5	21.9	370	2	Q92KX0	Q92KX0 helicobacte	927	37	21.6	255	1	Q9Y9Y5	Q9Y9Y5 aeropyrum p
855	37.5	21.9	372	5	O02305	O02305 caenorhabdi	928	37	21.6	259	2	Q53080	Q53080 rhodococcus
856	37.5	21.9	411	5	O17638	O17638 caenorhabdi	929	37	21.6	263	5	O61483	O61483 anopheles g
857	37.5	21.9	416	4	O15407	O15407 homo sapien	930	37	21.6	264	1	O28962	O28962 archaeoglob
858	37.5	21.9	423	5	O96090	O96090 ciona lites	931	37	21.6	264	2	O83326	O83326 treponema p
859	37.5	21.9	432	2	Q9ZDP3	Q9ZDP3 rickettsia	932	37	21.6	265	2	O68882	O68882 pseudomonas
860	37.5	21.9	432	10	Q9ZQW4	Q9ZQW4 populus tri	933	37	21.6	267	2	Q9X5B6	Q9X5B6 clostridium
861	37.5	21.9	448	2	P95531	P95531 pseudomonas	934	37	21.6	286	1	O26616	O26616 methanobact
862	37.5	21.9	449	4	O14570	O14570 homo sapien	935	37	21.6	286	2	P95553	P95553 pseudomonas
863	37.5	21.9	463	1	Q9YAG3	Q9YAG3 aeropyrum p	936	37	21.6	288	2	Q9X5B6	Q9X5B6 erwinia car
864	37.5	21.9	494	13	O57586	O57586 brachydanio	937	37	21.6	292	5	O44819	O44819 caenorhabdi
865	37.5	21.9	511	5	Q27900	Q27900 euplotes oc	938	37	21.6	293	2	O60203	O60203 plectonema
866	37.5	21.9	514	4	O75416	O75416 homo sapien	939	37	21.6	296	5	Q38098	Q38098 bacterioph
867	37.5	21.9	574	10	O82784	O82784 prunus sero	940	37	21.6	297	5	Q17417	Q17417 caenorhabdi
868	37.5	21.9	579	13	Q9Y100	Q9Y100 brachydanio	941	37	21.6	299	3	Q92NK8	Q92NK8 staphylococ
869	37.5	21.9	579	13	Q9W6E4	Q9W6E4 brachydanio	942	37	21.6	300	3	Q94457	Q94457 schizosacch
870	37.5	21.9	613	1	P81412	P81412 pyrococcus	943	37	21.6	301	2	O68183	O68183 enterococcu
871	37.5	21.9	657	11	O34732	O34732 mus musculu	944	37	21.6	303	4	Q15729	Q15729 homo sapien
872	37.5	21.9	682	5	O01327	O01327 caenorhabdi	945	37	21.6	308	1	O50782	O50782 methanobact
873	37.5	21.9	690	2	O30421	O30421 caldocellum	946	37	21.6	308	1	O27207	O27207 methanobact
874	37.5	21.9	701	5	P91922	P91922 calliphora	947	37	21.6	312	5	O16983	O16983 caenorhabdi
875	37.5	21.9	729	11	O35095	O35095 rattus norv	948	37	21.6	320	2	O68706	O68706 yersinia pe
876	37.5	21.9	729	11	Q9Z0E0	Q9Z0E0 mus musculu	949	37	21.6	322	5	O18406	O18406 drosophila
877	37.5	21.9	730	4	Q9Y4A6	Q9Y4A6 homo sapien	950	37	21.6	340	10	Q92NY7	Q92NY7 oryza sativ
878	37.5	21.9	731	4	Q9Y4D9	Q9Y4D9 homo sapien	951	37	21.6	345	2	O52130	O52130 escherichia
879	37.5	21.9	835	2	Q9X0A6	Q9X0A6 thermotoga	952	37	21.6	345	2	O85640	O85640 escherichia
880	37.5	21.9	1211	4	Q9S569	Q9S569 homo sapien	953	37	21.6	353	5	O54036	O54036 pseudomonas
881	37.5	21.9	1332	6	O97897	O97897 tragalaphus	954	37	21.6	354	5	O17701	O17701 caenorhabdi
882	37.5	21.9	1335	2	O33635	O33635 staphylococ	955	37	21.6	358	10	Q38877	Q38877 arabidopsis
883	37.5	21.9	1404	4	Q92954	Q92954 homo sapien	956	37	21.6	359	5	Q9XV41	Q9XV41 caenorhabdi
884	37.5	21.9	1487	5	O15843	O15843 leishmania	957	37	21.6	362	4	O95551	O95551 homo sapien
885	37.5	21.9	1516	4	P78354	P78354 homo sapien	958	37	21.6	365	2	O30835	O30835 rhodobacter
886	37.5	21.9	1532	4	P78544	P78544 homo sapien	959	37	21.6	366	1	O29817	O29817 archaeoglob
887	37.5	21.9	1707	12	Q96806	Q96806 grapevine t	960	37	21.6	366	2	O67371	O67371 aquifex aeo
888	37.5	21.9	1978	10	Q39849	Q39849 glycine max	961	37	21.6	366	10	O23825	O23825 pharbitis n
889	37.5	21.9	2322	4	Q92675	Q92675 homo sapien	962	37	21.6	367	10	Q96472	Q96472 ipomoea pur
890	37.5	21.9	2479	11	O63002	O63002 rattus norv	963	37	21.6	368	10	O92Q58	O92Q58 ipomoea bat
891	37	21.6	121	12	O73499	O73499 okra yellow	964	37	21.6	368	10	O92Q57	O92Q57 ipomoea bat
892	37	21.6	122	8	P92510	P92510 arabidopsis	965	37	21.6	370	2	O25558	O25558 helicobacte
893	37	21.6	127	1	Q9YEA5	Q9YEA5 aeropyrum p	966	37	21.6	371	1	O28901	O28901 archaeoglob
894	37	21.6	127	12	Q9Y211	Q9Y211 human immun	967	37	21.6	372	10	Q43262	Q43262 zea mays (m
895	37	21.6	127	12	Q9YYX6	Q9YYX6 human immun	968	37	21.6	374	4	Q9Y502	Q9Y502 homo sapien


```

969 37 21.6 390 1 028689
970 37 21.6 392 2 092501
971 37 21.6 399 5 027124
972 37 21.6 401 10 049403
973 37 21.6 401 10 082078
974 37 21.6 401 13 09W7E3
975 37 21.6 410 4 095244
976 37 21.6 412 2 052939
977 37 21.6 416 2 054238
978 37 21.6 416 2 055661
979 37 21.6 425 2 068581
980 37 21.6 430 5 026424
981 37 21.6 439 13 042464
982 37 21.6 441 3 013337
983 37 21.6 443 11 09WVJ9
984 37 21.6 445 2 050880
985 37 21.6 446 8 003639
986 37 21.6 450 3 059711
987 37 21.6 450 3 074829
988 37 21.6 452 11 09WVN9
989 37 21.6 453 2 066700
990 37 21.6 458 4 099952
991 37 21.6 459 2 069516
992 37 21.6 462 10 004622
993 37 21.6 462 12 039264
994 37 21.6 467 4 09Y286
995 37 21.6 467 13 042465
996 37 21.6 470 8 033508
997 37 21.6 470 8 033481
998 37 21.6 470 8 033483
999 37 21.6 472 1 058606

```

ALIGNMENTS

```

RESULT 1
Q15962 PRELIMINARY; PRT; 87 AA.
AC Q15962;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE THYROTROPIN BETA SUBUNIT (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 93103017.
RA "PEELEE M.E., CARR F.E., BAKER J.R. JR., WARTOFSKY L., BURMAN K.D.;
RT "TSH beta subunit gene expression in human lymphocytes.";
RL Am. J. Med. Sci. 305:1-7(1993).
DR EMBL; S51112; CAB30302.1; -.
DR HSSP; P01233; IHRP.
DR PFAM; PF00007; Cys_knot; 1.
FT NON_TER 1
SQ SEQUENCE 87 AA; 9911 MW; B09DF839 CRC32;

```

```

Query Match 95.9%; Score 164; DB 4; Length 87;
Best Local Similarity 96.7%; Pred. No. 1.6e-16;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 vctyrdfiyrtveipgcbllhvapyfsypva 30
|||||
DB 23 VCTYRDLHYKTFELPDCPLGVDPVTVPVA 52

```

```

RESULT 2
Q13051 PRELIMINARY; PRT; 150 AA.
ID Q13051
AC Q13051;

```

```

DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE THYROTROPIN BETA SUBUNIT PRECURSOR.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinidae; Cyprinidae; Cyprinus.
RN [1]
RP SEQUENCE FROM N.A.
RA KOBAYASHI M.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MAKINO K., KONDO H., IWASAKI M., YOSHIURA Y., WATABE S.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB003585; BAA20082.1; -.
DR HSSP; P01233; IHRP.
DR PFAM; PF00007; Cys_knot; 1.
DR PRINTS; PR00438; GFCYSKNOT.
KW Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 150 THYROTROPIN BETA SUBUNIT.
SQ SEQUENCE 150 AA; 17055 MW; 8122A6E8 CRC32;

```

```

Query Match 62.6%; Score 107; DB 13; Length 150;
Best Local Similarity 58.6%; Pred. No. 4.1e-08;
Matches 17; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 2 ctvyrdfiyrtveipgcbllhvapyfsypva 30
|||||
DB 71 CTYOEYRTALPFCPSHADPFTVPVA 99

```

```

RESULT 3
Q90225 PRELIMINARY; PRT; 137 AA.
ID Q90225
AC Q90225;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE GONADOTROPIN BETA-SUBUNIT.
OS Acanthopagrus latus (Yellowfin poiry).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Perciformes; Percoidae; Sparidae; Acanthopagrus.
RN [1]
RP SEQUENCE FROM N.A.
RA TSAI H.J., YANG L.T.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; L11722; AAA48512.1; -.
DR HSSP; P01233; IHRP.
DR PFAM; PF00007; Cys_knot; 1.
SQ SEQUENCE 137 AA; 15330 MW; F87577B5 CRC32;

```

```

Query Match 62.0%; Score 106; DB 13; Length 137;
Best Local Similarity 63.3%; Pred. No. 5.1e-08;
Matches 19; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 1 vctyrdfiyrtveipgcbllhvapyfsypva 30
|||||
DB 75 VCTYRDLHYKTFELPDCPLGVDPVTVPVA 104

```

```

RESULT 4
Q98849 PRELIMINARY; PRT; 140 AA.
ID Q98849
AC Q98849;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)

```


AC 013049;
DT 01-JUL-1997 (TREMBLrel. 04, Created)

GN	LHB.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-129 SVEV;
RA	KAPUR V., MATZUK M.M.;
RL	Submitted (APR-1995) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; U25145; AAA92841.1; --
DR	HSSP; P01233; IHRP.
DR	MGD; MGI:96782; LHD.
DR	PAM; PF00007; Cys_knot; 1.
DR	PRINTS; PR00438; GFCYSKNOT.
SQ	SEQUENCE 141 AA; 15028 MW; BE50AB2B CRC32;
Query Match	54.4%; Score 93; DB 11; Length 141;
Best Local Similarity	53.3%; Pred. No. 3.8e-06;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0	
QY	1 vctyrdfiyrtveipgcpvhvpyfsypva 30 : :
Db	76 VCTRELAFASVRLPGCPGVDPPIVSFPVA 105
RESULT 11	
Q62778	
ID	Q62778 PRELIMINARY; PRT; 139 AA.
AC	Q62778;
DT	01-NOV-1996 (TrEMBLrel. 01, Created)
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT	01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE	TESTICULAR LUTEINIZING HORMONE BETA-SUBUNIT.
GN	TLEH1.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-SPRAQUE DAWLEY;
RX	MEDLINE; 95283549.
RA	ZHANG F.P., RANNIKKO A., HUHTANIEMI I.;
RT	"Isolation and characterization of testis-specific cDNAs for
RT	luteinizing hormone beta-subunit in the rat."
RL	Biochem. Biophys. Res. Commun. 210:858-865(1995).
DR	EMBL; U25653; AAC52249.1; --
DR	HSSP; P01233; IHRP.
DR	PAM; PF00007; Cys_knot; 1.
DR	PRINTS; PR00438; GFCYSKNOT.
SQ	SEQUENCE 139 AA; 14880 MW; 5B2E7DBF CRC32;
Query Match	53.2%; Score 91; DB 11; Length 139;
Best Local Similarity	53.3%; Pred. No. 7.3e-06;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0	
QY	1 vctyrdfiyrtveipgcpvhvpyfsypva 30 : :
Db	74 VCTRELAFASVRLPGCPGVDPPIVSFPVA 103
RESULT 12	
O77805	
ID	O77805 PRELIMINARY; PRT; 143 AA.
AC	O77805;
DT	01-NOV-1998 (TrEMBLrel. 08, Created)
DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT	01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE	LUTEINIZING HORMONE BETA SUBUNIT PRECURSOR.
OS	Felis silvestris catus (Cat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC	Eutheria; Carnivora; Fissipedida; Felidae; Felis.


```

RN SEQUENCE FROM N.A.
RP TISSUE-PITUITARY GLAND;
RA PUKAZHENTHI B.S., VARMA G.M., BROWN J.L.;
RT "Molecular cloning and sequence analysis of the cDNA for the feline
RL luteinizing hormone beta subunit."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF095716; AAC64196.1;
DR HSSP: P01233; IHRP.
DR PFAM: PF00007; Cys_knot; 1.
DR PRINTS: PR00438; GFCYSKNOT.
SQ SEQUENCE 143 AA; 15318 MW; 52C3BABA CRC32;

Query Match 53.2%; Score 91; DB 6; Length 143;
Best Local Similarity 53.3%; Pred. No. 7.5e-06;
Matches 16; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpplhvafysypva 30
Db 78 VCTYRELRFASVRLPGCPGVDPVSPVA 107

RESULT 13
Q9W607 PRELIMINARY; PRT; 88 AA.
AC Q9W607;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE GONADOTROPIN II BETA SUBUNIT (FRAGMENT).
OS Salmo salar (Atlantic salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
OC Salmoniformes; Salmonidae; Salmo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PITUITARY;
RA YADETIE F., GOKSOYR A., MALE R.;
RT "Modulation of pituitary gonadotropin mRNA levels in juvenile Atlantic
RL salmon by 4-nonylphenol."
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF146151; AAD34593.1;
FT NON_TER 1
FT NON_TER 88
SQ SEQUENCE 88 AA; 9975 MW; 36FD1018 CRC32;

Query Match 52.6%; Score 90; DB 13; Length 88;
Best Local Similarity 53.3%; Pred. No. 6.2e-06;
Matches 16; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpplhvafysypva 30
Db 34 VCTYRDVRYETINLPDPPVWDHVTIPVA 63

RESULT 14
Q46618 PRELIMINARY; PRT; 89 AA.
AC Q46618;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE LUTEINIZING HORMONE/CHORIONIC GONADOTROPHIN BETA-SUBUNIT (FRAGMENT).
OS Equus zebra hartmannae.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Perissodactyla; Equidae; Equus.
RN [1]
RP SEQUENCE FROM N.A.
RA FISCHER S., VEITS J., MEYER H.H.D.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.

RN SEQUENCE 82 AA; 8643 MW; B6624078 CRC32;
```

```

DR EMBL: AF047602; AAC04360.1;
DR HSSP: P01233; IHRP.
DR PFAM: PF00007; Cys_knot; 1.
KW Chorion.
FT NON_TER 1
FT NON_TER 89
SQ SEQUENCE 89 AA; 9325 MW; B5989A39 CRC32;

Query Match 52.6%; Score 90; DB 6; Length 89;
Best Local Similarity 50.0%; Pred. No. 6.3e-06;
Matches 15; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpplhvafysypva 30
Db 15 VCTYRELRFASIRLPGCPGVDPVSPVA 44

RESULT 15
Q46619 PRELIMINARY; PRT; 89 AA.
AC Q46619;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE LUTEINIZING HORMONE/CHORIONIC GONADOTROPHIN BETA-SUBUNIT (FRAGMENT).
GN LH/CG-BETA.
OS Equus hemionus kulan (Kulan) (Asiatic wild ass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Perissodactyla; Equidae; Equus.
RN [1]
RP SEQUENCE FROM N.A.
RA FISCHER S., VEITS J., MEYER H.H.D.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF047603; AAC04361.1;
DR HSSP: P01233; IHRP.
DR PFAM: PF00007; Cys_knot; 1.
KW Chorion.
FT NON_TER 1
FT NON_TER 89
SQ SEQUENCE 89 AA; 9369 MW; B2136EE6 CRC32;

Query Match 52.6%; Score 90; DB 6; Length 89;
Best Local Similarity 50.0%; Pred. No. 6.3e-06;
Matches 15; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 vctyrdfiyrtveipgcpplhvafysypva 30
Db 16 VCTYRELRFASIRLPGCPGVDPVSPVA 45

RESULT 16
Q46622 PRELIMINARY; PRT; 82 AA.
AC Q46622;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE LUTEINIZING HORMONE BETA-SUBUNIT (FRAGMENT).
GN LH-BETA.
OS Ceratotherium simum simum.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Perissodactyla; Rhinocerotidae; Ceratotherium.
RN [1]
RP SEQUENCE FROM N.A.
RA FISCHER S., VEITS J., MEYER H.H.D.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF047607; AAC04365.1;
DR HSSP: P01233; IHRP.
DR PFAM: PF00007; Cys_knot; 1.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 8643 MW; B6624078 CRC32;
```



```

Query Match          49.7%; Score 85; DB 6; Length 82;
Best Local Similarity 46.7%; Pred. No. 3e-05;
Matches 14; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 vctyrdfllyrtveipgpcplhvafysypva 30
    |||| : : : |||| | | : ||||
Db 17 VCTYHFLFASIRLPGCPGVDPMVSPFA 46

RESULT 17
O19102          PRELIMINARY; PRT; 135 AA.
AC O19102;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE LUTEINIZING HORMONE BETA SUBUNIT (FRAGMENT).
OS Ceratotherium simum (White rhinoceros) (Square-lipped rhinoceros).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Perissodactyla; Rhinocerotidae; Ceratotherium.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PITUITARY GLAND;
RX MEDLINE: 97449288.
RA SHERMAN G.B., LUND L.A., BUNICK D., WINN R.J.;
RT "Characterization and phylogenetic significance of rhinoceros
RT luteinizing hormone beta (Lhbeta) subunit messenger RNA structure,
RT complementary DNA sequence and gene copy number.";
RL Gene 195:131-139(1997).
DR EMBL: U72659; AAB71983.1; -
DR HSSP: P01233; LHRP.
DR PFAM: PF000007; Cys_knot; 1.
FT NON_TER 1
SQ SEQUENCE 135 AA; 14212 MW; F4F196D2 CRC32;

Query Match          49.7%; Score 85; DB 6; Length 135;
Best Local Similarity 46.7%; Pred. No. 5.1e-05;
Matches 14; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 vctyrdfllyrtveipgpcplhvafysypva 30
    |||| : : : |||| | | : ||||
Db 70 VCTYHFLFASIRLPGCPGVDPMVSPFA 99

RESULT 18
O77835          PRELIMINARY; PRT; 141 AA.
AC O77835;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE LUTEINIZING HORMONE BETA SUBUNIT PRECURSOR.
GN LH BETA 2 OR LH BETA 1.
OS Ceratotherium simum (White rhinoceros) (Square-lipped rhinoceros).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Perissodactyla; Rhinocerotidae; Ceratotherium.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97449288.
RA SHERMAN G.B., LUND L.A., BUNICK D., WINN R.J.;
RT "Characterization and phylogenetic significance of rhinoceros
RT luteinizing hormone beta (Lhbeta) subunit messenger RNA structure,
RT complementary DNA sequence and gene copy number.";
RL Gene 195:131-139(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 98389253.
RA LUND L.A., SHERMAN G.B.;
RT "Duplication of the southern white rhinoceros (Ceratotherium simum
RT simum) luteinizing hormone beta subunit gene.";

```

```

J. Mol. Endocrinol. 21:19-30(1998).
DR EMBL: AF024521; AAC36049.1; -
DR EMBL: AF024520; AAC36048.1; -
DR HSSP: P01233; LHRP.
DR PFAM: PF00007; Cys_knot; 1.
DR PRINTS: PR00438; GFCYSKNOT.
KW SIGNAL.
FT SIGNAL.
FT CHAIN 1 20 POTENTIAL.
FT CHAIN 21 141 LUTEINIZING HORMONE BETA SUBUNIT.
SQ SEQUENCE 141 AA; 14930 MW; BAD6ED62 CRC32;

Query Match          49.7%; Score 85; DB 6; Length 141;
Best Local Similarity 46.7%; Pred. No. 5.4e-05;
Matches 14; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 vctyrdfllyrtveipgpcplhvafysypva 30
    |||| : : : |||| | | : ||||
Db 76 VCTYHFLFASIRLPGCPGVDPMVSPFA 105

RESULT 19
O9YIB3          PRELIMINARY; PRT; 130 AA.
AC O9YIB3;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE GONADOTROPIN I BETA SUBUNIT 2.
GN GFGTHIBETA-2.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinidae; Cyprinidae; Carassius.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 99051337.
RA SOHN Y.C., SUETAKE H., YOSHURA Y., KOBAYASHI M., AIDA K.;
RT "Structural and expression analyses of gonadotropin Ibeta subunit
RT genes in goldfish (Carassius auratus).";
RL Gene 222:257-267(1998).
DR EMBL: AB015483; BAA36975.1; -
DR HSSP: P01233; LHRP.
DR SEQUENCE 130 AA; 14449 MW; 62B2C1F5 CRC32;

Query Match          48.5%; Score 83; DB 13; Length 130;
Best Local Similarity 48.3%; Pred. No. 9.5e-05;
Matches 14; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 2 ctyrdfllyrtveipgpcplhvafysypva 30
    | | : | | | | | | | | | |
Db 74 CNFREWTYETFEKGPADSIFFSYPVA 102

RESULT 20
O9YGH2          PRELIMINARY; PRT; 149 AA.
AC O9YGH2;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE GONADOTROPIN II-BETA PRECURSOR.
GN GTH II-BETA.
OS Clupea harengus (Atlantic herring).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Clupeomorpha; Clupeidae; Clupea.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PITUITARY;
RA POWER M.E., CAROLSFIELD J., WALLIS G.P., SHERWOOD N.M.;
RT "Isolation and characterization of a cDNA for gonadotropin II-beta of
RT Pacific herring, an ancient teleost.";

```



```

RN SEQUENCE FROM N.A.
RP TISSUE-PITUITARY;
RA VADETIE P., GOKSOYR A., MALE R.;
RT "Modulation of pituitary gonadotropin mRNA levels in juvenile Atlantic
RL salmon by 4-nonylphenol.";
DR EMBL: AF146152; AAD34594.1; -
FT NON_TER 1 1
FT NON_TER 123 123
SQ SEQUENCE 123 AA; 13656 MW; 4990D5E7 CRC32;

Query Match 39.5%; Score 67.5; DB 13; Length 123;
Best Local Similarity 44.8%; Pred. No. 0.015;
Matches 13; Conservative 5; Mismatches 10; Indels 1; Gaps 1;

Qy 2 cttyrdfiyrtveipgcp--lhvapyfsypva 30
   |:::| |:::| |:::| |:::|
Db 66 CNFKDWSYKVKYLEGCPGSGVDPPF-IPVA 93

RESULT 25
ID O67006 PRELIMINARY; PRT; 262 AA.
AC O67006;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE FLAGELLAR HOOK BASAL-BODY PROTEIN FLGG.
GN FLGG2.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RX MEDLINE; 98196666.
RA GRAHAM D.E., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RN Nature 392:353-358(1998).
RL [2]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE000709; AAC06963.1; -
DR PFAM: PF00460; flg_bb_rod; 1.
KW Flagella.
SQ SEQUENCE 262 AA; 28360 MW; 9CAF857F CRC32;

Query Match 32.7%; Score 56; DB 2; Length 262;
Best Local Similarity 56.2%; Pred. No. 1.5;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 3 tyrdfiyrtveipgcp 18
   |:::| |:::| |:::| |:::|
Db 40 TFQDLYQTVTEPGAP 55

RESULT 26
ID Q66148 PRELIMINARY; PRT; 108 AA.
AC Q66148;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE ORF1, PROTEINASE, RNA-DEPENDENT RNA POLYMERASE, AND COAT PROTEIN.

```

```

OS cocksfoot mottle virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Sobemovirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RUSSIA;
RA RYABOV E.V., KRUTOV A.A., NOVIKOV V.K., ZELENKOVA O.V.,
RA MOROZOV S.YU., ZAVRIEV S.K.;
RL Phytopathology 0:0-0(0).
DR EMBL: L40905; AAA87608.1; -
SQ SEQUENCE 108 AA; 12348 MW; 2BD7FFE7 CRC32;

Query Match 28.7%; Score 49; DB 12; Length 108;
Best Local Similarity 41.9%; Pred. No. 5.8;
Matches 13; Conservative 4; Mismatches 8; Indels 6; Gaps 2;

Qy 1 vctyrdfiyrtveipgcp--lhvapyfsypv 29
   |:::| |:::| |:::| |:::|
Db 70 ICEFRE----TVEIPSLPKGFKVSSDFSYSV 96

RESULT 27
ID Q91120 PRELIMINARY; PRT; 120 AA.
AC Q91120;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE GONADOTROPIN I BETA SUBUNIT PRECURSOR.
OS Morone saxatilis (Striped bass).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorpha;
OC Perciformes; Percoidae; Moronidae; Morone.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PITUITARY GLAND;
RX MEDLINE; 96020549.
RA HASSIN S., ELIZUR A., ZOHAR Y.;
RT "Molecular cloning and sequence analysis of striped bass (Morone
RT saxatilis) gonadotrophin-I and -II subunits.";
RL J. Mol. Endocrinol. 15:23-35(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-PITUITARY GLAND;
RA HASSIN S.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: L35070; AAC38035.1; -
DR HSP; P01233; 1HRP.
DR PFAM: PF00007; Cys_knot; 1.
KW Signal.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 120 GONADOTROPIN I BETA SUBUNIT.
SQ SEQUENCE 120 AA; 13220 MW; 1F55A817 CRC32;

Query Match 28.7%; Score 49; DB 13; Length 120;
Best Local Similarity 44.0%; Pred. No. 6.5;
Matches 11; Conservative 3; Mismatches 7; Indels 4; Gaps 1;

Qy 6 dfiyrtveipgcp--lhvapyfsypva 30
   |:::| |:::| |:::| |:::|
Db 72 DWSYEVKHKICPGV-----TYPVA 92

RESULT 28
ID O48730 PRELIMINARY; PRT; 292 AA.
AC O48730;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE T9022.30 PROTEIN.
GN T9022.30.

```


OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC eukaryotes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
 OC Arabidopsids.
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=CV. COLOMBIA;
 RA ROUNSLEY S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C.,
 RA STOKES S.M., KAUL S., MASON T.M., KERLAVAGE A.R., ADAMS M.D.,
 RA SOMERVILLE C.R., VENTER J.C.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC002505; AAC14510.1; -;
 SQ SEQUENCE 292 AA; 33801 MW; B5EA8CA7 CRC32;

Query Match 28.7%; Score 49; DB 10; Length 292;
 Best Local Similarity 33.3%; Pred. No. 17;
 Matches 13; Conservative 5; Mismatches 9; Indels 12; Gaps 3;

QY 1 vcty-rdflyrtveipg-----cplhvapysfyp 28
 :|: ||| | :|: | :|: ||| :|:
 Db 79 ICNFDKFIYAYGVPGRAHDTKVLNYCATN-EPYFSHP 116

RESULT 29
 ID P94612 PRELIMINARY; PRT; 452 AA.
 AC P94612;

DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last annotation update)
 DE ORF452.
 OS Coxiella burnetii.
 OC Bacteria; Proteobacteria; gamma subdivision; Coxiella.
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=NINE MILE RSA493 PHASE I;
 RA WILLEMS H., JAEGER C.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Y10436; CAA71458.1; -;
 SQ SEQUENCE 452 AA; 49730 MW; FCD39824 CRC32;

Query Match 28.7%; Score 49; DB 2; Length 452;
 Best Local Similarity 47.4%; Pred. No. 27;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 tyrdflyrtveipgclhv 21
 :|: ||| :|: | :|: ||| :|:
 Db 250 TTRDIRESIHIDGLPIHV 268

RESULT 30
 ID Q63315 PRELIMINARY; PRT; 813 AA.
 AC Q63315;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
 DE LONG TYPE PB-CADHERIN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=WISTAR; TISSUE=BRAIN;
 RX MEDLINE; 96212232.
 RA SUGIMOTO K., HONDA S., YAMAMOTO T., UEKI T., MONDEN M., KAJI A.,
 RA MATSUMOTO K., NAKAMURA T.;
 RT "Molecular cloning and characterization of a newly identified member
 of the cadherin family, PB-cadherin.";
 RL J. Biol. Chem. 271:11548-11556(1996).

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 DR EMBL: D83348; BAA11894.1; -;
 DR HSP: P15116; INCJ
 DR PROSITE; PS00232; CADHERIN; 2.
 DR PFAM; PF00028; cadherin; 5.
 DR PFAM; PF01049; cadherin_C-term; 1.
 DR PRINTS; PR00205; CADHERIN.
 KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat.
 SQ SEQUENCE 813 AA; 87978 MW; 6EAD9F1D CRC32;

Query Match 28.7%; Score 49; DB 11; Length 813;
 Best Local Similarity 41.7%; Pred. No. 50;
 Matches 10; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 4 yrdflyrtveipgclhvapysfyp 27
 :|: ||| | :|: | :|: ||| :|:
 Db 732 FRDFISRKVALADADLSVPPYDAF 755

RESULT 31
 ID Q9WTP5 PRELIMINARY; PRT; 813 AA.
 AC Q9WTP5;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
 DE PB-CADHERIN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=ICR;
 RA KITAJIMA K., KOSHIMIZU K., NAKAMURA T.;
 RT "Molecular cloning and characterization of a newly identified member
 of the cadherin family, PB-cadherin.";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=ICR;
 RX MEDLINE; 96212232.
 RA SUGIMOTO K., HONDA S., YAMAMOTO T., UEKI T., MONDEN M., KAJI A.,
 RA MATSUMOTO K., NAKAMURA T.;
 RT "Molecular cloning and characterization of a newly identified member
 of the cadherin family, PB-cadherin.";
 RL J. Biol. Chem. 271:11548-11556(1996).
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
 DR EMBL: AB019618; BAA34426.1; -;
 DR PROSITE; PS00232; CADHERIN; 2.
 KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat.
 SQ SEQUENCE 813 AA; 88021 MW; 5CEDA1CC CRC32;

Query Match 28.7%; Score 49; DB 11; Length 813;
 Best Local Similarity 41.7%; Pred. No. 50;
 Matches 10; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 4 yrdflyrtveipgclhvapysfyp 27
 :|: ||| | :|: | :|: ||| :|:
 Db 732 FRDFISRKVALADADLSVPPYDAF 755

RESULT 32
 ID Q50041 PRELIMINARY; PRT; 47 AA.
 AC Q50041;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE REGX.
 OS Mycobacterium leprae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;


```

OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RA SMITH D.R.;
RA Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RA ROBISON K.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U15182; AAA62997.1; -.
DR PFAM; PF00486; trans_reg_C; 1.
SQ SEQUENCE 47 AA; 5263 MW; 90E19575 CRC32;

Query Match 28.4%; Score 48.5; DB 2; Length 47;
Best Local Similarity 69.2%; Pred. No. 2.8;
Matches 9; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

Qy 8 iyrvtveipgcplh 20
   |||:|:|:|:|
Db 4 IV-TARLPGCPLH 15

RESULT 33
Q18351
ID Q18351 PRELIMINARY; PRT; 491 AA.
AC Q18351;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
DE C32C4.1 PROTEIN.
GN C32C4.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA MCMURRAY A.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPOAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL; 273905; CRA98109.1; -.
DR PFAM; PF00520; ion_trans; 1.
DR PRINTS; PR00169; KCHANNEL.
SQ SEQUENCE 491 AA; 56324 MW; 5962F19A CRC32;

Query Match 28.4%; Score 48.5; DB 5; Length 491;
Best Local Similarity 31.2%; Pred. No. 34;
Matches 10; Conservative 5; Mismatches 10; Indels 7; Gaps 1;

Qy 5 rdfivrtveipgcplhvapy-----fsypv 29
   :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 246 KDFVYKPTDSPNLPULILEYICIGWTFEYLV 277

RESULT 34
Q26497
ID Q26497 PRELIMINARY; PRT; 148 AA.

```

```

AC Q26497;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE FORMATE HYDROGENLYASE, SUBUNIT 7.
GN MTH397.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DELTA H;
RX MEDLINE; 98037514.
RA SMITH D.R., DOUCETTE-STAMM L.A., DELOUGHERY C., LEE H.-M., DUBOIS J.,
RA ALDREDGE T., BASHIRZADEH R., BLAKELY D., COOK R., GILBERT K.,
RA HARRISON D., HOANG L., KEAGLE P., LUMM W., POTHIER B., QIU D.,
RA SPADAFORA R., VICARE R., WANG Y., WIERZBOWSKI J., GIBSON R.,
RA JIWANI N., CARUSO A., BUSH D., SAFER H., PATWELL D., PRABHAKAR S.,
RA MCDUGALL S., SHIMER G., GOYAL A., PIETROVSKI S., CHURCH G.M.,
RA DANIELS C.J., MAO J.-I., RICE P., NOLLING J., REEVE J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7153(1997).
DR EMBL; AE000824; AAB84903.1; -.
DR PFAM; PF01058; oxidored_g6; 1.
KW Lyase.
SQ SEQUENCE 148 AA; 16079 MW; A4E9B511 CRC32;

Query Match 28.1%; Score 48; DB 1; Length 148;
Best Local Similarity 42.9%; Pred. No. 11;
Matches 12; Conservative 1; Mismatches 5; Indels 10; Gaps 1;

Qy 6 dfivrtveipgcplh-----lhvap 23
   |||:|:|:|:|:|:|:|:|:|:|:|
Db 111 DFIPVDVPGCPRPSPSEILEAILAVAP 138

RESULT 35
Q49138
ID Q49138 PRELIMINARY; PRT; 307 AA.
AC Q49138;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE (CLONE PDN9, HINDIIIAB).
GN MXAA.
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Methylobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AMI;
RX MEDLINE; 96074311.
RA MORRIS C.J., KIM Y.M., PERKINS K.E., LIDSTROM M.E.;
RT "Identification and nucleotide sequences of mxaA, mxaC, mxaL,
RT and mxaD genes from Methylobacterium extorquens AM1.";
RL J. Bacteriol. 177:6825-6831(1995).
DR EMBL; L41608; AAA85568.1; -.
KW Signal.
RT SIGNAL.
FT CHAIN 1 19 POTENTIAL.
FT CHAIN 20 307
SQ SEQUENCE 307 AA; 33649 MW; ECFB79E2 CRC32;

Query Match 28.1%; Score 48; DB 2; Length 307;
Best Local Similarity 43.5%; Pred. No. 24;
Matches 10; Conservative 5; Mismatches 4; Indels 4; Gaps 1;

Qy 3 tyrdfiy-----rtveipgcplhv 21
   |||:|:|:|:|:|:|:|:|:|:|:|
Db 93 TYQDFYVALDARTLDVPGFPVTV 115

```


RESULT 36
Q9ZIS9 PRELIMINARY; PRT; 327 AA.
AC Q9ZIS9;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE PUTATIVE BETA1,3-GLUCOSYLTRANSFERASE WAAV.
GN WAAV.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-F470;
RX MEDLINE; 99009057.
RA HEINRICH D.E., YETHON J.A., AMOR P.A., WHITFIELD C.;
RT "The assembly system for the outer core portion of R1- and R4-type
RT lipopolysaccharides of Escherichia coli. The R1 core-specific beta-
RT glucosyltransferase provides a novel attachment site for O-
RT polysaccharides.";
RL J. Biol. Chem. 273:29497-29505(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-F470;
RX MEDLINE; 98434532.
RA YETHON J.A., HEINRICH D.E., MONTEIRO M.A., PERRY M.B., WHITFIELD C.;
RT "Involvement of waaY, waaQ, and waaP in the modification of
RT Escherichia coli lipopolysaccharide and their role in the formation of
RT a stable outer membrane.";
RL J. Biol. Chem. 273:26310-26316(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-F470;
RX MEDLINE; 99009352.
RA HEINRICH D.E., YETHON J.A., WHITFIELD C.;
RT "Molecular basis for structural diversity in the core regions of the
RT lipopolysaccharides of Escherichia coli and Salmonella enterica.";
RL Mol. Microbiol. 30:221-232(1998).
DR EMBL; AF019746; AAC69672.1; -.
KW Transferase.
SQ SEQUENCE 327 AA; 38779 MW; AF588FE4 CRC32;

Query Match 28.1%; Score 48; DB 2; Length 327;
Best Local Similarity 40.9%; Pred. No. 26;
Matches 9; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

QY 6 dfiyrvtveipgclhvapyfsy 27
||||| : | : | : | : |
DB 186 dfiyrL--LITCDMYAPVYIM 205

RESULT 37
Q03289 PRELIMINARY; PRT; 374 AA.
AC Q03289;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
DE D8035.3P.
GN TOM1 OR D8035.3.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RA DIETRICH F.S., MULLIGAN J., ALLEN E., ARAUJO R., AVILES E., BERNO A.,
RA CARPENTER J., CHEN E., CHERRY J.M., CHUNG E., DUNCAN M.,
RA HUNICKE-SMITH S., HYMAN R., KOMP C., LASHKARI D., LEW H., LIN D.,
RA MOSDALE D., NAKAHARA K., NAMATH A., OEFNER P., OH C., PETEL F.X.,
RA ROBERTS D., SCHRAMM S., SCHROEDER M., SHOGREN T., SHROFF N.,
RA WINANT A., YELTON M., BOTSTEIN D., DAVIS R.W.;

RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA DIETRICH F.S.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA JIA Y., CHERRY J.M.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U33030; AAB64921.1; -.
DR SGD; L0002983; TOM1.
DR PFAM; PF01529; zf-DHHC; 1.
SQ SEQUENCE 374 AA; 42922 MW; 9BE81442.CRC32;

Query Match 27.8%; Score 47.5; DB 3; Length 374;
Best Local Similarity 40.7%; Pred. No. 36;
Matches 11; Conservative 4; Mismatches 9; Indels 3; Gaps 1;

QY 7 fiyrvtvei---pgclhvapyfsypva 30
||| : | : | : | : | : |
DB 69 fiwLQIVILVPGTQPHVAPFLILPIA 95

RESULT 38
O88306 PRELIMINARY; PRT; 320 AA.
AC O88306;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
DE HUPS HOMOLOG.
OS Nostoc PCC73102.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PCC73102 (ATCC 29133);
RX MEDLINE; 98200551.
RA OXELFELT F., TAMAGNINI P., LINDBLAD P.;
RT "Hydrogen uptake in Nostoc sp. strain PCC 73102. Cloning and
RT characterization of a hupSL homologue.";
RL Arch. Microbiol. 169:267-274(1998).
DR EMBL; AF030525; AAC16276.1; -.
DR HSSP; P21853; 1H2A.
SQ SEQUENCE 320 AA; 34919 MW; 98226828.CRC32;

Query Match 27.5%; Score 47; DB 2; Length 320;
Best Local Similarity 40.0%; Pred. No. 36;
Matches 8; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

QY 5 rdfiyr-----tveipgclph 20
||| : | : | : | : |
DB 145 KDFVSQAGLPVINIPGCPAH 164

RESULT 39
O88036 PRELIMINARY; PRT; 359 AA.
AC O88036;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
DE OXIDOREDUCTASE ALPHA SUBUNIT.
GN SC5A7.30C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA SEEGER K.J., HARRIS D.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.


```

RN  SEQUENCE FROM N.A.
RP  STRAIN=A3(2);
RA  PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL  Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN  [3]
RP  SEQUENCE FROM N.A.
RX  MEDLINE: 97000351.
RA  KINASHI H., HOPWOOD D.A.;
RT  "A set of ordered cosmids and a detailed genetic and physical map for
RI  the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL  Mol. Microbiol. 21:77-96(1996).
DR  EMBL: AL031107; CAAL1995.1; -.
DR  PFAM: PF00355; Rieske; I.
SQ  SEQUENCE 359 AA; 39575 MW; 34595B78 CRC32;

Query Match      27.5%; Score 47; DB 2; Length 359;
Best Local Similarity 43.2%; Pred. No. 40;
Matches 16; Conservative 1; Mismatches 12; Indels 8; Gaps 3;

QY  1 vctyrdfyrt----veipg---cplhvapyfsypva 30
    ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  73 VCGYHGFTYDTGTCVYVPGKRPV-RTARVASYPVA 108

RESULT 40
O53282 PRELIMINARY; PRT; 360 AA.
ID AC O53282;
DT 01-JUN-1998 (T-EMBLrel. 06, Created)
DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)
DT 01-MAY-1999 (T-EMBLrel. 10, Last annotation update)
DE HYPOTHETICAL 37.3 KD PROTEIN.
GN MTV012.50.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RA OLIVER K., HARRIS D.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA BARRELL B.G., RAJANDREAM M.A., PARKHILL J., COLE S.T.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE: 96181548.
RA PHILIPP W.J., POULET S., EIGLMEIER K., PASCOPELLA L.,
RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,
RA COLE S.T.;
RT "An integrated map of the genome of the tubercle bacillus,
RT Mycobacterium tuberculosis H37RV, and comparison with Mycobacterium
RT leprae.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:3132-3137(1996).
DR EMBL: AL021287; CAAL1620.1; -.
DR PFAM: PF01011; Bacterial_PQO; 1.
KW Hypothetical protein.
SQ SEQUENCE 360 AA; 37323 MW; 816824D4 CRC32;

Query Match      27.5%; Score 47; DB 2; Length 360;
Best Local Similarity 66.7%; Pred. No. 40;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY  15 pgcplhvapyfs 26
    ||||| ||| |||
```

```

Db  144 PGCPVAAAPAFS 155

RESULT 41
O01838 PRELIMINARY; PRT; 383 AA.
ID AC O01838;
DT 01-JUL-1997 (T-EMBLrel. 04, Created)
DT 01-JUL-1997 (T-EMBLrel. 04, Last sequence update)
DT 01-NOV-1998 (T-EMBLrel. 08, Last annotation update)
DE COSMID B0414.
GN B0414.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., McMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SHALDON N., SMITH A., SONNHAMMER E., STADEN R., SOLSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA SAMMONS L., WOHLDMANN P., ROHLFING T.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA WATERSTON R.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003145; AAB57714.1; -.
SQ SEQUENCE 383 AA; 43840 MW; 17B6C588 CRC32;

Query Match      27.5%; Score 47; DB 5; Length 383;
Best Local Similarity 28.6%; Pred. No. 43;
Matches 12; Conservative 8; Mismatches 10; Indels 12; Gaps 1;

QY  1 vctyrdfyrtvtel-----pgcplhvapyfsypva 30
    ||| :||| ||| ||| ||| ||| ||| ||| ||| |||
Db  66 ICSMQFYIYVTLSIGAITWLFDFDLTAAGCVFIVSLFFAYFA 107

RESULT 42
O67030 PRELIMINARY; PRT; 448 AA.
ID AC O67030;
DT 01-AUG-1998 (T-EMBLrel. 07, Created)
DT 01-AUG-1998 (T-EMBLrel. 07, Last sequence update)
DT 01-NOV-1998 (T-EMBLrel. 08, Last annotation update)
DE THIOPHENE AND FURAN OXIDATION PROTEIN.
GN THDF.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE: 98196666.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBEEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
```


Query Match 27.2%; Score 46.5; DB 5; Length 334;
Best Local Similarity 39.3%; Pred. NO. 44;
Matches 11; Conservative 4; Mismatches 12; Indels 1; Gaps 1;

Qy	2	ctyrdfiyrtveipgcplhvapýfsypv	29
	1:	1::111:	11111111
Db	6	CSITDYLX-TPDFFETVLHVISIPIV	32

RESULT 46

049857
ID 049857 PRELIMINARY; PRT; 341 AA.
AC 049857;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE ASPARAGINE-BOND SPECIFIC CYSTEINE ENDOPEPTASE.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
OC Glycine.
RN [1]
RP SEQUENCE FROM N.A.
RA FUKAZAWA C.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
DR ENBL; X78548; CAA5294.1; -;
DR MENDEL; 27772; Glyma.1147;27772.
DR PFAM; PF00561; abhydrolase; 1.
DR PRINTS; PR00412; EPOXYDLASE.
KW Protease.
SQ SEQUENCE 341 AA; 39170 MW; 57B97259 CRC32;

Query Match 26.9%; Score 46; DB 10; Length 341;
Best Local Similarity 47.6%; Pred. No. 53;
Matches 10; Conservative 3; Mismatches 8; Indels

Qy 9 yrtveipgcp1hvapyfsypv 29
:1111:1:111111
Db 31 HRTVEVNGIKMHVAEKGECPV 51

RESULT 47

[illegible]

Query Match 26.9%; Score 46; DB 10; Length 341;
Best Local Similarity 47.6%; Pred. NO. 53;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 9 yrtveipgcp1hvapyfsypv 29
:1111:1:111111
Db 31 HRTVEVNGIKMHVAEKGECPV 51

RESULT 48

043205	PRELIMINARY;	PRT;	380 AA.
ID	043205		
AC	01-JUN-1998	(TREMBLrel. 06, Created)	
AD	01-JUN-1998	(TREMBLrel. 06, Last sequence update)	
DT	01-JUN-1998	(TREMBLrel. 06, Last sequence update)	
DE	01-NOV-1999	(TREMBLrel. 12, Last annotation update)	
DE	CLONE 23663	(FRAGMENT).	
OS	Homo sapiens	(Human).	
OC	Eukaryota;	Metazoa;	Chordata;
OC	Eutheria;	Primates;	Catarrhini;
OC		Hominidae;	Homo.
RI	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=BRAIN;		
RX	ADLSSINE; 96207227.		
RA	ANDERSSON B., WENTLAND M.A., RICAFFRENTE J.Y., LIU W., GIBBS R.A.;		
RT	"A 'double adaptor' method for improved shotgun library		
RT	construction."		
RL	Anal. Biochem. 236:107-113(1996).		
RP	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=BRAIN;		
RX	MEDLINE; 97264341.		
RA	YU W., ANDERSSON B., WORLEY K.C., MUZYNY D.M., DING Y., LIU W.,		
RA	RICAFFRENTE J.Y., WENTLAND M.A., LENNON G., GIBBS R.A.;		
RT	"Large-scale concatenation cDNA sequencing."		
RT	Genome Res. 7:353-358(1997).		
CC	-I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).		
DR	EMBL; AF035300; AAB88183.1; -.		
DR	PROSITE; PS00232; CADHERIN; 1.		
DR	PFAM; PF00028; cadherin; 1.		
DR	PFAM; PF01049; cadherin_Cterm; 1.		
KW	Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat.		
FT	NON_TER		
FT	1		
SO	SEQUENCE	380 AA; 40254 MW; 0E1C2BE0 CRC32;	

Query Match 26.9%; Score 46; DB 4; Length 380;
Best Local Similarity 41.7%; Pred. NO. 59;
Matches 10; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

```
Qy      4 yrdfiyrtveipgcplhvapyfsy 27
      :||| | | : | | | :
Db     299 FRDFISRKVALADGDLSPVPYDAF 32
```

RESULT 49

AC	074497	PRELIMINARY;	PRT;	382 AA.
ID	074497			
AC	074497;			
DT	01-NOV-1998	(TREMBLrel. 08, Created)		
DT	01-NOV-1998	(TREMBLrel. 08, Last sequence update)		
DT	01-MAY-1999	(TREMBLrel. 10, Last annotation update)		
DE	HYPOTHETICAL RYANODINE RECEPTOR DOMAIN CONTAINING PROTEIN.			
GN	SPC2B5.10C.			
OS	Schizosaccharomyces pombe (Fission yeast).			
OC	Eukaryota; Fungi; Ascomycota; Archiascomycetes;			
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;			
OC	Schizosaccharomycetes.			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN=972H-;			

RA SIEGER K., HARRIS D., LYNE M., RAJANDREAM M.A., BARRELL B.G.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031545; CAA20849.1; -
DR PFAM; PF00622; SPRY; 1.
SQ SEQUENCE 382 AA; 42869 MW; 2C1CFC3A CRC32;

Query Match 26.9%; Score 46; DB 3; Length 382;
Best Local Similarity 57.1%; Pred. No. 60;
Matches 12; Conservative 2; Mismatches 5; Indels 2; Gaps 2;

Qy	2	ctyrdfiyrtveipg-cplhv	21
		:	
Db	258	CTYRN-LYPTVGAIGPCTLHV	277

RESULT 50

ID	005950	PRELIMINARY;	PRT;	711 AA.
AC	005950;			
DT	01-JUL-1997	(TrEMBLrel. 04, Created)		
DT	01-AUG-1999	(TrEMBLrel. 11, Last sequence update)		
DT	01-AUG-1999	(TrEMBLrel. 11, Last annotation update)		
DE	HYPOTHETICAL 82.1 KD PROTEIN.			
GN	RP674.			
OS	Rickettsia prowazekii.			
OC	Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;			
OC	Rickettsiaceae; Rickettsia;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	STRAIN=MADRID E;			
RC	STRAIN=MADRID E;			
RX	MEDLINE: 99039499.			
RA	ANDERSSON S.G.E.; ZOMORODIPOUR A.; ANDERSSON J.O.;			
RA	SICHERITZ-PONTEN T.; ALSMARK U.C.M.; PODOWSKI R.M.; NAEGLUND A.K.;			
RA	ERIKSSON A.S.; WINKLER H.H.; KURLAND C.G.;			
RT	"The genome sequence of Rickettsia prowazekii and the origin of mitochondria."			
RT	Nature 396:133-140(1998).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RP	STRAIN=MADRID E;			
RC	STRAIN=MADRID E;			
RA	ANDERSSON S.G.E.;			
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE OF 111-711 FROM N.A.			
RP	STRAIN=MADRID E;			
RC	STRAIN=MADRID E;			
RX	MEDLINE: 97419517.			
RA	ANDERSSON J.O.; ANDERSSON S.G.E.;			
RT	"Genomic rearrangements during evolution of the obligate intracellular parasite Rickettsia prowazekii as inferred from an analysis of 520 bp nucleotide sequence."			
RT	Microbiology 143:2783-2795(1997).			
RL	EMBL; AJ235272; CAAL511.1;			
RL	EMBL; Y11786; CAAT2480.1;			
DR	Hypothetical protein.			
SW	SEQUENCE 711 AA; 82138 MW; 3E3B72BA CRC32;			

Query Match 26.9%; Score 46; DB 2; Length 711;
Best Local Similarity 39.1%; Pred. No. 1.2e+02;
Matches 9: Conservative 4; Mismatches 8; Indels 2; Gaps 1;

Qy 8 l y r t v - - e l p g c p l h v a p y f s y p 28
: : : | | | | : | | | : |
Db 171 V Y K N V F E E Y E G V P I F Y L P Y F F H P 193

Search completed: July 14, 2000, 09:35:35
Job time: 2386 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 14, 2000, 09:31:00 ; Search time 37.7 Seconds
(without alignments)
16.335 Million cell u

Title: ALPHA-CHAIN
perfect score: 147
Sequence: 1 apdvqdcpctlganpffsqgqapil 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 188963 seqs, 23686106 residues

Total number of hits satisfying chosen parameters:

```
Minimum DB seq length: 0
Maximum DB seq length: 1000000
```

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 1000 summaries

Database : A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			ID	Description
	Score	Match	Length		
1	139	94.6	116	1	W9544
2	138	93.9	116	1	hCG analogue alpha
3	137	93.2	116	1	hCG analogue alpha
4	137	93.2	116	1	hCG analogue alpha
5	121	82.3	68	1	hCG/bCG alpha subu
6	121	82.3	96	1	hCG/bCG alpha subu
7	121	82.3	96	1	hCG/bCG alpha subu
8	121	82.3	96	1	hCG/bCG alpha subu
9	115	78.2	256	1	TBP(20-161)/hCG-al
10	115	78.2	285	1	TBP(20-190)/hCG-al
11	111	75.5	92	1	hCG/bCG alpha subu
12	110	74.8	92	1	hCG/bCG alpha subu
13	106	72.1	96	1	hCG/bCG alpha subu
14	106	72.1	96	1	hCG/bCG alpha subu
15	106	72.1	96	1	hCG/bCG alpha subu
16	105	71.4	96	1	hCG/bCG alpha subu
17	100	68.0	96	1	hCG/bCG alpha subu
18	87	59.2	15	1	Synthetic glycopro
19	85	57.8	96	1	hCG/bCG alpha subu
20	85	57.8	96	1	hCG/bCG alpha subu
21	85	57.8	120	1	Ovine FSH alpha su
22	81	55.1	15	1	Synthetic glycopro
23	78	53.1	120	1	Equine alpha subun
24	78	53.1	120	1	Equine chorionic g
25	78	53.1	120	1	Equine chorionic g
26	78	53.1	120	1	Equine chorionic g
27	78	53.1	120	1	Equine chorionic g
28	78	53.1	120	1	Equine chorionic g
29	78	53.1	265	1	Equine chorionic g
30	72	49.0	89	1	Salmon gonadotroph
31	72	49.0	119	1	Fish gonadotrophic
32	65	44.2	12	1	Human TSH alpha-su
33	65	44.2	12	1	Human TSH alpha-su

107	41	27.9	138	1	R10097	Engineered partial	180	38.5	26.2	154	1	W35850	Human CD7 for use
108	41	27.9	454	1	R88904	Human insulin rece	181	38.5	26.2	350	1	R05385	Xylene oxygenase g
109	41	27.9	457	1	W97843	Rat bone mineralis	182	38.5	26.2	437	1	P94413	Sequence of coding
110	41	27.9	457	1	W97844	Human bone mineral	183	38.5	26.2	483	1	R79144	Human tissue PA va
111	41	27.9	552	1	R30775	PH52-8.0 humanised	184	38.5	26.2	483	1	R70884	Human tissue PA va
112	41	27.9	787	1	W37912	Homo sapiens Smoot	185	38.5	26.2	483	1	R70885	Human tissue PA va
113	41	27.9	787	1	W81064	Amino acid sequenc	186	38.5	26.2	483	1	R70886	Human tissue PA va
114	41	27.9	839	1	R75109	Glycosyl-phosphati	187	38.5	26.2	483	1	R70887	Human tissue PA va
115	40.5	27.6	1013	1	W40223	Murine m11l protei	188	38.5	26.2	483	1	R70888	Human tissue PA va
116	40.5	27.6	1139	1	W37779	Rattus norvegicus	189	38.5	26.2	483	1	R70850	Human tissue PA va
117	40.5	27.6	1251	1	W37778	Rattus norvegicus	190	38.5	26.2	483	1	R70852	Human tissue PA va
118	40.5	27.6	1323	1	R55248	N-methyl-D-asparti	191	38.5	26.2	483	1	R70851	Human tissue PA va
119	40.5	27.6	1336	1	R66041	Human N-methyl-D-a	192	38.5	26.2	483	1	R70853	Human tissue PA va
120	40.5	27.6	1336	1	W87511	Human N-methyl-D-a	193	38.5	26.2	483	1	R70854	Human tissue PA va
121	40.5	27.6	1911	1	R99534	Dermatomyositis sp	194	38.5	26.2	483	1	R70855	Human tissue PA va
122	40	27.2	77	1	W28214	Staphylococcus aur	195	38.5	26.2	650	1	P70084	Sequence of hybrid
123	40	27.2	118	1	R12497	Human 5' EST secre	196	38.5	26.2	682	1	R49144	Product of alterna
124	40	27.2	145	1	R15092	hCG/HTSH chimera,	197	38.5	26.2	682	1	R87154	Alternatively spli
125	40	27.2	216	1	W37716	C. glutamicum orf3	198	38.5	26.2	708	1	Y13355	Amino acid sequenc
126	40	27.2	289	1	W70806	A human inorganic	199	38.5	26.2	836	1	R58912	Product of alterna
127	40	27.2	289	1	W67881	Human secreted pro	200	38.5	26.2	836	1	R87153	Alternatively spli
128	40	27.2	381	1	Y05323	Human secreted pro	201	38.5	26.2	904	1	R58907	Human protocadheri
129	40	27.2	464	1	R20020	t-PADeltaFAFE/u-PA-e	202	38.5	26.2	904	1	R87147	Protocadherin clon
130	40	27.2	464	1	R20021	t-PADeltaFAFEK1delta	203	38	25.9	20	1	W13739	Mouse bax protein
131	40	27.2	548	1	R20019	t-PADeltaK1K2/u-PA	204	38	25.9	20	1	W53625	Bax peptide (aa43-
132	40	27.2	548	1	R20018	t-PA/u-PA-e. New c	205	38	25.9	24	1	R84539	Anti-HIV-1 gp160 v
133	40	27.2	587	1	W51314	Human semaphorin W	206	38	25.9	96	1	R41279	ISG1 ORF-1 prod.
134	40	27.2	592	1	R92527	Fas antigen #2. Im	207	38	25.9	143	1	R20718	C10-35 NANBH-speci
135	40	27.2	671	1	R94953	Type I (beta-1) ra	208	38	25.9	169	1	R78612	Plasmid fragment p
136	40	27.2	673	1	P80697	Sequence of rat br	209	38	25.9	209	1	R20612	C10-14 NANBH-speci
137	40	27.2	673	1	R94764	Type II (beta-2) r	210	38	25.9	233	1	Y07075	Renal cancer assoc
138	40	27.2	776	1	W51313	Rat semaphorin W.	211	38	25.9	251	1	W74968	Human secreted pro
139	40	27.2	1021	1	W23281	Alloreaction assoc	212	38	25.9	255	1	R20611	C10-13 NANBH-speci
140	40	27.2	4303	1	R90302	Polycystic kidney	213	38	25.9	263	1	W87769	Human tissue plas
141	39.5	26.9	376	1	R36773	Human fibromodulin	214	38	25.9	263	1	Y05219	Kringel protein s
142	39.5	26.9	376	1	W26404	Human fibromodulin	215	38	25.9	327	1	R41688	Murine Fas antigen
143	39.5	26.9	449	1	W19720	Human smooth muscl	216	38	25.9	327	1	R78611	Murine Fas antigen
144	39.5	26.9	449	1	W56151	Procollagen C-prot	217	38	25.9	327	1	R92530	mFas sequence. Imm
145	39.5	26.9	507	1	W42333	W382C mutant of in	218	38	25.9	327	1	W86241	Fas ligand (FasL)
146	39.5	26.9	524	1	R98523	DnaA gene product.	219	38	25.9	342	1	R66213	Nocardia corallina
147	39.5	26.9	619	1	W75771	Human GTP binding	220	38	25.9	343	1	R81469	Nocardia corallina
148	39.5	26.9	1833	1	R79478	Mouse LTBSP-2. Tran	221	38	25.9	362	1	R81100	Sequence of RI fra
149	39	26.5	102	1	W78211	Human secreted pro	222	38	25.9	380	1	R50075	NANBH virus antige
150	39	26.5	297	1	R43575	Human adrenocortic	223	38	25.9	426	1	R71583	Raphanus sativus f
151	39	26.5	297	1	W19707	Melanocortin-2 rec	224	38	25.9	436	1	R74803	Corynebacterium gl
152	39	26.5	297	1	W79685	Melanocortin-2 rec	225	38	25.9	436	1	R92519	Corynebacterium gl
153	39	26.5	297	1	W92440	Human MC2 protein.	226	38	25.9	475	1	R20613	C10-15 NANBH-speci
154	39	26.5	344	1	W70335	Staphylococcus aur	227	38	25.9	477	1	R74175	Human collapsin. N
155	39	26.5	373	1	W60853	Bovine CD14 protei	228	38	25.9	485	1	W32426	Mycobacterium tube
156	39	26.5	406	1	W05321	Human mucosal addr	229	38	25.9	485	1	W32358	Mycobacterium tube
157	39	26.5	406	1	Y02068	Human MadCAM-1 pro	230	38	25.9	485	1	W64238	Mycobacterium tube
158	39	26.5	515	1	R70995	Human/rat alpha-1B	231	38	25.9	485	1	W81661	M. tuberculosis im
159	39	26.5	515	1	R90041	Human/rat hybrid a	232	38	25.9	498	1	W19748	Drosophila inhibit
160	39	26.5	515	1	W77106	Rat alpha18-adrene	233	38	25.9	546	1	W40115	Human alpha-5(IV)
161	39	26.5	520	1	R53072	Alpha 1b adrenergi	234	38	25.9	576	1	R78613	Expression vector
162	39	26.5	520	1	R52831	Sequence of human	235	38	25.9	596	1	R39554	Deduced amino acid
163	39	26.5	520	1	R85943	Alpha-1B adrenergi	236	38	25.9	615	1	R31349	Jagdslekte retrovi
164	39	26.5	587	1	R99256	Natural killer lyt	237	38	25.9	685	1	R89115	Bombay mori (pro)p
165	39	26.5	620	1	W14994	Human c-Fos induce	238	38	25.9	685	1	W14441	Prophenol oxidase.
166	39	26.5	737	1	Y05221	Human Stat6b prote	239	38	25.9	771	1	R71380	Human semaphorin I
167	39	26.5	793	1	W37911	Rattus norvegicus	240	38	25.9	1211	1	R67629	Non-A Non-B hepati
168	39	26.5	793	1	W81065	Amino acid sequenc	241	38	25.9	1263	1	W22052	DNA polymerase 385
169	39	26.5	819	1	Y05222	Human Stat6c prote	242	38	25.9	1263	1	W97098	Pfu DNA polymerase
170	39	26.5	841	1	R22376	Human liver GPI-PL	243	38	25.9	2050	1	W73499	Von Willebrand fac
171	39	26.5	841	1	R75110	Glycosyl-phosphati	244	38	25.9	2050	1	W90117	Human mature von W
172	39	26.5	847	1	R88320	IL-4 Stat peptide.	245	38	25.9	2476	1	W67738	Pig p105 zona pell
173	39	26.5	848	1	Y05223	Human Stat6b protei	246	38	25.9	2594	1	W14748	IgG-Fc binding pro
174	39	26.5	1435	1	R70232	P. falciparum SABP	247	38	25.9	2813	1	P60053	Sequence of von W
175	39	26.5	1435	1	W22477	Silicic acid bindin	248	38	25.9	2813	1	P60462	Sequence of human
176	39	26.5	1528	1	R33363	Human multidrug re	249	38	25.9	2864	1	R82072	Hepatitis GB virus
177	39	26.5	1604	1	R70105	TNF-R-EBA 175 fusi	250	38	25.9	3033	1	R33539	NANBH virus strain
178	39	26.5	1786	1	R41043	CD4-EBA175 fusion	251	38	25.9	3033	1	R33214	NANBH virus strain
179	39	26.5	3084	1	W50891	Mouse laminin A ch	252	38	25.9	3033	1	R33538	NANBH virus strain

253	38	25.9	5405	1	W14749	IgG-Fc binding pro	326	37.5	25.5	483	1	R09248	t-PA variant dl-44
254	37.5	25.5	157	1	P94405	Native tPA gene pr	327	37.5	25.5	483	1	R09249	t-PA variant dl-44
255	37.5	25.5	224	1	R11494	Endothelin-3 precu	328	37.5	25.5	483	1	R09250	t-PA variant dl-44
256	37.5	25.5	224	1	R25159	Endothelin-3 precu	329	37.5	25.5	483	1	R09251	t-PA variant dl-44
257	37.5	25.5	238	1	R23785	Prepro ET-3 sequen	330	37.5	25.5	483	1	R09252	t-PA variant dl-44
258	37.5	25.5	285	1	W14465	Baboon acrosomal s	331	37.5	25.5	483	1	R09253	t-PA variant dl-44
259	37.5	25.5	285	1	W14466	Macaque acrosomal s	332	37.5	25.5	483	1	R09254	t-PA variant dl-44
260	37.5	25.5	310	1	P80770	Recombinant human	333	37.5	25.5	483	1	R09257	t-PA variant dl-44
261	37.5	25.5	348	1	R35401	tPA-2. Tissue plas	334	37.5	25.5	483	1	R09263	t-PA variant Y67N,
262	37.5	25.5	348	1	R35404	tPA-8. Tissue plas	335	37.5	25.5	483	1	R09263	Human tissue PA va
263	37.5	25.5	355	1	P94407	Sequence encoded b	336	37.5	25.5	483	1	R70877	Human tissue PA va
264	37.5	25.5	355	1	P94411	Sequence of coding	337	37.5	25.5	483	1	R70878	Human tissue PA va
265	37.5	25.5	355	1	R06417	Tissue plasminogen	338	37.5	25.5	483	1	R70879	Human tissue PA va
266	37.5	25.5	355	1	R35400	tPA-1. Tissue plas	339	37.5	25.5	483	1	R70880	Human tissue PA va
267	37.5	25.5	355	1	R35402	tPA-6. Tissue plas	340	37.5	25.5	483	1	R70881	Human tissue PA va
268	37.5	25.5	355	1	R35403	tPA-7. Tissue plas	341	37.5	25.5	483	1	R70882	Human tissue PA va
269	37.5	25.5	355	1	R35405	tPA-11. Tissue pla	342	37.5	25.5	483	1	R70883	Human tissue PA va
270	37.5	25.5	355	1	R35406	tPA-12. Tissue pla	343	37.5	25.5	483	1	R70890	Human tissue PA va
271	37.5	25.5	355	1	R35407	tPA-13. Tissue pla	344	37.5	25.5	483	1	R70843	Human tissue PA va
272	37.5	25.5	355	1	R35408	tPA-14. Tissue pla	345	37.5	25.5	483	1	R70844	Human tissue PA va
273	37.5	25.5	355	1	R35409	tPA-15. Tissue pla	346	37.5	25.5	483	1	R70845	Human tissue PA va
274	37.5	25.5	355	1	R35410	tPA-16. Tissue pla	347	37.5	25.5	483	1	R70846	Human tissue PA va
275	37.5	25.5	355	1	R35411	tPA-17. Tissue pla	348	37.5	25.5	483	1	R70847	Human tissue PA va
276	37.5	25.5	355	1	R35412	tPA-18. Tissue pla	349	37.5	25.5	483	1	R70848	Human tissue PA va
277	37.5	25.5	355	1	R35413	tPA-5. Tissue plas	350	37.5	25.5	483	1	R70849	Human tissue PA va
278	37.5	25.5	355	1	R35414	tPA-6. Tissue plas	351	37.5	25.5	516	1	P70257	Sequence of human
279	37.5	25.5	355	1	R44835	Human truncated tp	352	37.5	25.5	516	1	P60614	Plasmd pdAP3 enco
280	37.5	25.5	355	1	R44836	Tissue plasminogen	353	37.5	25.5	518	1	P91961	Sequence of des 1-
281	37.5	25.5	355	1	W14431	mt-PA6-E. Recombin	354	37.5	25.5	518	1	R09237	t-PA deletion vari
282	37.5	25.5	355	1	W28536	Thrombolytically a	355	37.5	25.5	518	1	R74681	t-PA mutein (N117Q
283	37.5	25.5	355	1	W34667	Tissue plasminogen	356	37.5	25.5	518	1	R74684	t-PA mutein (N117Q
284	37.5	25.5	356	1	W14430	Tissue plasminogen	357	37.5	25.5	518	1	R74686	t-PA mutein (N117Q
285	37.5	25.5	356	1	W14432	Tissue plasminogen	358	37.5	25.5	518	1	R74679	t-PA mutein (N117Q
286	37.5	25.5	356	1	W28535	mt-PA9-D. Recombin	359	37.5	25.5	519	1	R09236	t-PA deletion vari
287	37.5	25.5	356	1	W28537	[GARSYQ]-[Plasmino	360	37.5	25.5	519	1	R74687	t-PA mutein (N117Q
288	37.5	25.5	371	1	R22502	Human tissue plasm	361	37.5	25.5	519	1	R74680	t-PA mutein (N117Q
289	37.5	25.5	372	1	W87770	rchd523 gene produ	362	37.5	25.5	519	1	R68852	Delta 55-62 tissue
290	37.5	25.5	375	1	W03739	FEQ-1 protein. DNA	363	37.5	25.5	520	1	R09235	t-PA deletion vari
291	37.5	25.5	375	1	W22338	Human rchd523 prot	364	37.5	25.5	520	1	R74688	t-PA mutein (N117Q
292	37.5	25.5	375	1	W89302	Amino acid sequenc	365	37.5	25.5	520	1	R74689	t-PA mutein (N117Q
293	37.5	25.5	375	1	W95158	Sequence of coding	366	37.5	25.5	521	1	P50400	tPA-(13-527) . Synt
294	37.5	25.5	375	1	P94417	Modified tissue pl	367	37.5	25.5	521	1	R09234	t-PA deletion vari
295	37.5	25.5	380	1	P82588	Sequence of coding	368	37.5	25.5	521	1	R09242	t-PA deletion vari
296	37.5	25.5	393	1	P70021	Recombinant tissue	369	37.5	25.5	521	1	R74682	t-PA mutein (N117Q
297	37.5	25.5	393	1	P70019	mt-PA6. New digly	370	37.5	25.5	521	1	R74678	t-PA mutein (N117Q
298	37.5	25.5	396	1	R25190	Chain 1 of modifie	371	37.5	25.5	522	1	R13917	Delta (466-470) tp
299	37.5	25.5	433	1	P81914	Delta(92-179) and	372	37.5	25.5	522	1	R13918	Delta (466-470) tp
300	37.5	25.5	434	1	R13922	Sequence of coding	373	37.5	25.5	522	1	R13919	Delta (466-470) tp
301	37.5	25.5	437	1	P94412	Sequence of coding	374	37.5	25.5	522	1	R13920	Delta (466-470) tp
302	37.5	25.5	437	1	P94410	Sequence of coding	375	37.5	25.5	522	1	R13921	Delta (466-470) tp
303	37.5	25.5	438	1	P94409	t-PA variant d92-1	376	37.5	25.5	522	1	R14486	t-PA deletion vari
304	37.5	25.5	439	1	R09258	t-PA variant d92-1	377	37.5	25.5	522	1	R09233	t-PA deletion vari
305	37.5	25.5	439	1	R09259	t-PA variant d92-1	378	37.5	25.5	522	1	R09241	t-PA deletion vari
306	37.5	25.5	439	1	R09260	t-PA variant d92-1	379	37.5	25.5	522	1	R74683	t-PA mutein (N117Q
307	37.5	25.5	439	1	R09261	Delta 2-89 tissue	380	37.5	25.5	522	1	R74685	t-PA mutein (N117Q
308	37.5	25.5	439	1	R68851	Modified tissue pl	381	37.5	25.5	522	1	W32113	Tumour necrosis fa
309	37.5	25.5	441	1	P70643	Sequence of tissue	382	37.5	25.5	523	1	R09232	t-PA deletion vari
310	37.5	25.5	445	1	R09285	Non-glycosylated t	383	37.5	25.5	523	1	R09240	t-PA deletion vari
311	37.5	25.5	446	1	R08150	[GARSYQ]-[Plasmino	384	37.5	25.5	524	1	P82584	Modified tissue pl
312	37.5	25.5	467	1	R22499	Sequence of coding	385	37.5	25.5	524	1	P82585	t-PA deletion vari
313	37.5	25.5	472	1	P94416	Modified tissue pl	386	37.5	25.5	524	1	R09231	t-PA deletion vari
314	37.5	25.5	473	1	P70642	Tissue plasminogen	387	37.5	25.5	524	1	R09239	Mutant human tissu
315	37.5	25.5	476	1	R30004	Mutant alpha-amyla	388	37.5	25.5	524	1	R99131	Mutant human tissu
316	37.5	25.5	478	1	R46065	Aspergillus oryzae	389	37.5	25.5	524	1	R99132	Mutant human tissu
317	37.5	25.5	478	1	R72450	Aspergillus oryzae	390	37.5	25.5	524	1	R99133	Mutant human tissu
318	37.5	25.5	478	1	R78270	Mature taka-amylas	391	37.5	25.5	524	1	R99134	Mutant human tissu
319	37.5	25.5	478	1	R79025	Aspergillus oryzae	392	37.5	25.5	524	1	R99135	Mutant human tissu
320	37.5	25.5	478	1	W14500	Modified tissue pl	393	37.5	25.5	524	1	R99136	Mutant human tissu
321	37.5	25.5	479	1	P82583	t-PA variant M3100	394	37.5	25.5	524	1	R99137	Mutant human tissu
322	37.5	25.5	480	1	R22714	Polypeptide with p	395	37.5	25.5	524	1	R99138	Mutant human tissu
323	37.5	25.5	482	1	P60613	t-PA variant dl-44	396	37.5	25.5	525	1	P90175	Tissue plasminogen
324	37.5	25.5	483	1	R09246	t-PA variant dl-44	397	37.5	25.5	525	1	P93540	t-PA deletion vari
325	37.5	25.5	483	1	R09247	t-PA variant dl-44	398	37.5	25.5	525	1	R09230	t-PA deletion vari

399	37.5	25.5	525	1	R09238	t-PA deletion vari	472	37.5	25.5	527	1	R09267	t-PA variant D283A
400	37.5	25.5	525	1	R09243	t-PA deletion vari	473	37.5	25.5	527	1	R09268	t-PA variant K296A
401	37.5	25.5	525	1	R09244	t-PA deletion vari	474	37.5	25.5	527	1	R09269	t-PA variant E303A
402	37.5	25.5	525	1	R09245	t-PA deletion vari	475	37.5	25.5	527	1	R09270	t-PA variant H331A
403	37.5	25.5	526	1	R09221	t-PA deletion vari	476	37.5	25.5	527	1	R09271	t-PA variant R339A
404	37.5	25.5	526	1	R09222	t-PA deletion vari	477	37.5	25.5	527	1	R09272	t-PA variant E347A
405	37.5	25.5	526	1	R09223	t-PA deletion vari	478	37.5	25.5	527	1	R09273	t-PA variant D364A
406	37.5	25.5	526	1	R09224	t-PA deletion vari	479	37.5	25.5	527	1	R44808	Human tPA variant
407	37.5	25.5	526	1	R09225	t-PA deletion vari	480	37.5	25.5	527	1	R44809	Human tPA variant
408	37.5	25.5	526	1	R09226	t-PA deletion vari	481	37.5	25.5	527	1	R44810	Human tPA variant
409	37.5	25.5	526	1	R09227	t-PA deletion vari	482	37.5	25.5	527	1	R44811	Human tPA variant
410	37.5	25.5	526	1	R09228	t-PA deletion vari	483	37.5	25.5	527	1	R44812	Human tPA variant
411	37.5	25.5	526	1	R09229	t-PA deletion vari	484	37.5	25.5	527	1	R44813	Human tPA variant
412	37.5	25.5	527	1	R04186	Plasminogen activa	485	37.5	25.5	527	1	R44814	Human tPA variant
413	37.5	25.5	527	1	R05489	PAO24 precursor p	486	37.5	25.5	527	1	R44815	Human tPA variant
414	37.5	25.5	527	1	R05488	tPA024 precursor p	487	37.5	25.5	527	1	R44816	Human tPA variant
415	37.5	25.5	527	1	R05806	Thrombolytic prote	488	37.5	25.5	527	1	R44817	Human tPA variant
416	37.5	25.5	527	1	R06236	Novel tissue plasm	489	37.5	25.5	527	1	R70866	Human t-PA variant
417	37.5	25.5	527	1	R22621	Mutated recombinan	490	37.5	25.5	527	1	R70867	Human t-PA variant
418	37.5	25.5	527	1	R22620	Mutated recombinan	491	37.5	25.5	527	1	R70868	Human t-PA variant
419	37.5	25.5	527	1	R22619	Mutated recombinan	492	37.5	25.5	527	1	R70869	Human t-PA variant
420	37.5	25.5	527	1	R22618	Mutated recombinan	493	37.5	25.5	527	1	R70870	Human t-PA variant
421	37.5	25.5	527	1	R22617	Mutated recombinan	494	37.5	25.5	527	1	R70871	Human t-PA variant
422	37.5	25.5	527	1	P71747	Gln(117)-substitut	495	37.5	25.5	527	1	R70872	Human t-PA variant
423	37.5	25.5	527	1	P71748	Tissue plasminogen	496	37.5	25.5	527	1	R70873	Human t-PA variant
424	37.5	25.5	527	1	P71749	Tissue plasminogen	497	37.5	25.5	527	1	R70874	Human t-PA variant
425	37.5	25.5	527	1	P71750	Tissue plasminogen	498	37.5	25.5	527	1	R70875	Human t-PA variant
426	37.5	25.5	527	1	P70168	Sequence of wild-t	499	37.5	25.5	527	1	R70876	Human t-PA variant
427	37.5	25.5	527	1	P70880	Sequence of thromb	500	37.5	25.5	527	1	R70898	Human t-PA variant
428	37.5	25.5	527	1	R11801	Tissue plasminogen	501	37.5	25.5	527	1	R70899	Human t-PA variant
429	37.5	25.5	527	1	R11343	Tissue plasminogen	502	37.5	25.5	527	1	R70900	Human t-PA variant
430	37.5	25.5	527	1	R11342	Tissue plasminogen	503	37.5	25.5	527	1	R70901	Human t-PA variant
431	37.5	25.5	527	1	R11345	Tissue plasminogen	504	37.5	25.5	527	1	R70902	Human t-PA variant
432	37.5	25.5	527	1	R11346	Tissue plasminogen	505	37.5	25.5	527	1	R70903	Human t-PA variant
433	37.5	25.5	527	1	R11347	Tissue plasminogen	506	37.5	25.5	527	1	R70904	Human t-PA variant
434	37.5	25.5	527	1	R13910	T-PA deriv. (I). N	507	37.5	25.5	527	1	R70905	Human t-PA variant
435	37.5	25.5	527	1	R13911	T-PA deriv. (II).	508	37.5	25.5	527	1	R70906	Human t-PA variant
436	37.5	25.5	527	1	R13912	T-PA deriv. (III).	509	37.5	25.5	527	1	R70907	Human t-PA variant
437	37.5	25.5	527	1	R13913	T-PA deriv. (IV).	510	37.5	25.5	527	1	R70908	Human t-PA variant
438	37.5	25.5	527	1	R13914	T-PA deriv. (V). N	511	37.5	25.5	527	1	R70909	Human t-PA variant
439	37.5	25.5	527	1	P91683	Sequence of tissue	512	37.5	25.5	527	1	R70893	Human t-PA variant
440	37.5	25.5	527	1	R20217	t-PA analogue expr	513	37.5	25.5	527	1	R70894	Human t-PA variant
441	37.5	25.5	527	1	R20223	t-PA analogue expr	514	37.5	25.5	527	1	R70895	Human t-PA variant
442	37.5	25.5	527	1	R20222	t-PA analogue expr	515	37.5	25.5	527	1	R70896	Human t-PA variant
443	37.5	25.5	527	1	R20221	t-PA analogue expr	516	37.5	25.5	527	1	R70897	Human t-PA variant
444	37.5	25.5	527	1	R20220	t-PA analogue expr	517	37.5	25.5	527	1	R70842	Wild type tissue p
445	37.5	25.5	527	1	R20215	R462E t-PA analogu	518	37.5	25.5	527	1	R70857	Human t-PA variant
446	37.5	25.5	527	1	R20216	R462G t-PA analogu	519	37.5	25.5	527	1	R70858	Human t-PA variant
447	37.5	25.5	527	1	R20218	t-PA analogue expr	520	37.5	25.5	527	1	R70859	Human t-PA variant
448	37.5	25.5	527	1	R20219	t-PA analogue expr	521	37.5	25.5	527	1	R70860	Human t-PA variant
449	37.5	25.5	527	1	R21592	tPA variant - gene	522	37.5	25.5	527	1	R70861	Human t-PA variant
450	37.5	25.5	527	1	R21593	tPA variant - E94A	523	37.5	25.5	527	1	R70862	Human t-PA variant
451	37.5	25.5	527	1	R21594	tPA variant - D95A	524	37.5	25.5	527	1	R70863	Human t-PA variant
452	37.5	25.5	527	1	R21595	tPA variant - D95G	525	37.5	25.5	527	1	R70864	Human t-PA variant
453	37.5	25.5	527	1	R21596	tPA variant - E94A	526	37.5	25.5	527	1	R70865	Human t-PA variant
454	37.5	25.5	527	1	R21597	tPA variant - E94A	527	37.5	25.5	527	1	R99124	Mutant human tissu
455	37.5	25.5	527	1	R21598	tPA variant - T103	528	37.5	25.5	527	1	R99125	Mutant human tissu
456	37.5	25.5	527	1	R21599	tPA variant - N117	529	37.5	25.5	527	1	R99126	Mutant human tissu
457	37.5	25.5	527	1	R21600	tPA variant - E94A	530	37.5	25.5	527	1	R99127	Mutant human tissu
458	37.5	25.5	527	1	R09274	t-PA variant E408A	531	37.5	25.5	527	1	R99128	Mutant human tissu
459	37.5	25.5	527	1	R09275	t-PA variant E410A	532	37.5	25.5	527	1	R99129	Mutant human tissu
460	37.5	25.5	527	1	R09276	t-PA variant K416A	533	37.5	25.5	527	1	R99130	Mutant human tissu
461	37.5	25.5	527	1	R09277	t-PA variant E426A	534	37.5	25.5	527	1	W45907	Single chain form
462	37.5	25.5	527	1	R09278	t-PA variant H432A	535	37.5	25.5	527	1	W57778	R275E,H417D human
463	37.5	25.5	527	1	R09279	t-PA variant R440A	536	37.5	25.5	527	1	W57779	R275E,H417E human
464	37.5	25.5	527	1	R09280	t-PA variant H445A	537	37.5	25.5	527	1	W57780	R275E,H429Y human
465	37.5	25.5	527	1	R09281	t-PA variant R449A	538	37.5	25.5	528	1	P90171	Tissue plasminogen
466	37.5	25.5	527	1	R09282	t-PA variant D460A	539	37.5	25.5	528	1	R70733	Thrombolytic prote
467	37.5	25.5	527	1	R09283	t-PA variant D477A	540	37.5	25.5	528	1	P71451	Modified human tis
468	37.5	25.5	527	1	R09215	t-PA variant F305	541	37.5	25.5	528	1	P71449	Modified human tis
469	37.5	25.5	527	1	R09255	t-PA variant I210R	542	37.5	25.5	528	1	P71450	Modified human tis
470	37.5	25.5	527	1	R09256	t-PA variant I210R	543	37.5	25.5	528	1	R25435	T-PA variant R299D
471	37.5	25.5	527	1	R09262	t-PA variant Y67N,	544	37.5	25.5	528	1	R09216	t-PA insertion var

545	37.5	25.5	528	1	R09218	t-PA insertion var	618	37.5	25.5	562	1	P50342	Human tPA. Vector
546	37.5	25.5	528	1	R68853	Bifibronection doma	619	37.5	25.5	562	1	P70285	Sequence encoded b
547	37.5	25.5	529	1	R09217	t-PA insertion var	620	37.5	25.5	562	1	P30001	Sequence of full l
548	37.5	25.5	529	1	R09219	t-PA insertion var	621	37.5	25.5	562	1	R23802	Zymogen-like t-PA
549	37.5	25.5	529	1	R09220	t-PA insertion var	622	37.5	25.5	562	1	R23801	Zymogen-like t-PA
550	37.5	25.5	530	1	P70449	Sequence encoded b	623	37.5	25.5	562	1	R23803	t-PA (Ser 304) mut
551	37.5	25.5	530	1	P70879	Sequence of thromb	624	37.5	25.5	562	1	R23804	t-PA (Glu 304) mut
552	37.5	25.5	530	1	P70881	Sequence of thromb	625	37.5	25.5	562	1	R23806	t-PA (Glu 296) mut
553	37.5	25.5	530	1	P70882	Sequence of thromb	626	37.5	25.5	562	1	R23807	t-PA (Tyr 297) mut
554	37.5	25.5	530	1	P71659	Sequence of thromb	627	37.5	25.5	562	1	R23808	t-PA (Glu 298) mut
555	37.5	25.5	530	1	P92277	Sequence of modifi	628	37.5	25.5	562	1	R23809	t-PA (Glu 299) mut
556	37.5	25.5	530	1	P61007	Sequence of tissue	629	37.5	25.5	562	1	R23810	t-PA (Gly 301) mut
557	37.5	25.5	531	1	R07034	Thrombolytic prote	630	37.5	25.5	562	1	R23811	t-PA (Glu 296, Glu
558	37.5	25.5	531	1	P60700	Sequence of tissue	631	37.5	25.5	562	1	R09284	Sequence of tissue
559	37.5	25.5	537	1	R11661	Tissue plasminogen	632	37.5	25.5	562	1	R09286	Sequence of tissue
560	37.5	25.5	539	1	R12846	T-PA growth domain	633	37.5	25.5	562	1	R09287	Sequence of tissue
561	37.5	25.5	549	1	P71745	des(Cys51-Asp87)t-	634	37.5	25.5	562	1	R09288	Sequence of tissue
562	37.5	25.5	552	1	R11662	Tissue plasminogen	635	37.5	25.5	562	1	R09289	Sequence of tissue
563	37.5	25.5	555	1	R06238	Novel mutant tissu	636	37.5	25.5	562	1	R09290	Sequence of tissue
564	37.5	25.5	555	1	R23805	t-PA (Del 296-302)	637	37.5	25.5	562	1	R35399	TPA7. Tissue plasm
565	37.5	25.5	556	1	R13148	T-PA variant contg	638	37.5	25.5	562	1	R34426	Sequence of human
566	37.5	25.5	557	1	R13020	T-PA variant contg	639	37.5	25.5	562	1	R38674	Sequence of tissue
567	37.5	25.5	557	1	R13149	T-PA variant contg	640	37.5	25.5	562	1	R41545	Modified tPA MB101
568	37.5	25.5	558	1	R13150	T-PA with -ve char	641	37.5	25.5	562	1	R44834	Human tPA (R129W) .
569	37.5	25.5	558	1	R13153	T-PA with -ve char	642	37.5	25.5	562	1	R96220	Full-length tissue
570	37.5	25.5	559	1	P94402	Analogue of amino	643	37.5	25.5	562	1	W47535	Tissue plasminogen
571	37.5	25.5	559	1	P81361	Tissue plasminogen	644	37.5	25.5	562	1	W47536	Tissue plasminogen
572	37.5	25.5	559	1	R06054	Tissue plasminogen	645	37.5	25.5	562	1	W47537	Tissue plasminogen
573	37.5	25.5	559	1	R12340	T-PA variant contg	646	37.5	25.5	570	1	R07477	t-PA variant MB102
574	37.5	25.5	559	1	R13151	T-PA with -ve char	647	37.5	25.5	571	1	P90169	Tissue plasminogen
575	37.5	25.5	559	1	R13152	T-PA with -ve char	648	37.5	25.5	580	1	P90168	Tissue plasminogen
576	37.5	25.5	559	1	R13154	T-PA variant havin	649	37.5	25.5	586	1	P71141	t-PA with extra fi
577	37.5	25.5	559	1	R13155	T-PA variant havin	650	37.5	25.5	586	1	P71144	des(Cys51-Asp87)t-
578	37.5	25.5	559	1	R13156	T-PA variant havin	651	37.5	25.5	592	1	P90172	Tissue plasminogen
579	37.5	25.5	559	1	R96223	Novel plasminogen	652	37.5	25.5	593	1	P70059	Sequence of hybrid
580	37.5	25.5	559	1	R96222	Novel plasminogen	653	37.5	25.5	594	1	P71743	Ile(277)t-PA with
581	37.5	25.5	559	1	W23368	Human tissue plasm	654	37.5	25.5	605	1	R06239	Novel mutant tissu
582	37.5	25.5	560	1	R12341	T-PA variant contg	655	37.5	25.5	623	1	P71140	t-PA with extra fi
583	37.5	25.5	561	1	P70020	Sequence of tissue	656	37.5	25.5	623	1	P71142	Ile(277)t-PA with
584	37.5	25.5	561	1	R12342	T-PA with -ve char	657	37.5	25.5	623	1	W72197	HSV-2 strain SB5 C
585	37.5	25.5	561	1	R12367	T-PA with -ve char	658	37.5	25.5	650	1	P70060	Sequence of hybrid
586	37.5	25.5	562	1	P90916	Human tissue plasm	659	37.5	25.5	680	1	P90173	Tissue plasminogen
587	37.5	25.5	562	1	P93716	Human melanoma t-p	660	37.5	25.5	704	1	P90176	Tissue plasminogen
588	37.5	25.5	562	1	P94238	Human tissue plasm	661	37.5	25.5	718	1	P90177	Tissue plasminogen
589	37.5	25.5	562	1	P94378	Amino acid sequenc	662	37.5	25.5	737	1	R76640	Deltex protein. No
590	37.5	25.5	562	1	P94379	Amino acid sequenc	663	37.5	25.5	737	1	R76639	Deltex protein pro
591	37.5	25.5	562	1	P94380	Amino acid sequenc	664	37.5	25.5	737	1	W18317	Drosophila Deltex
592	37.5	25.5	562	1	P94406	Sequence encoded b	665	37.5	25.5	779	1	P90178	Tissue plasminogen
593	37.5	25.5	562	1	R03388	Thrombolytic prote	666	37.5	25.5	780	1	R06824	Thrombomodulin ana
594	37.5	25.5	562	1	R04700	Sequence of tissue	667	37.5	25.5	794	1	P90179	Tissue plasminogen
595	37.5	25.5	562	1	R04699	Native tissue plas	668	37.5	25.5	807	1	R20013	PA mutant Plg 1-54
596	37.5	25.5	562	1	R04702	Sequence of tissue	669	37.5	25.5	1039	1	P80692	Hybrid plasminogen
597	37.5	25.5	562	1	R04701	Sequence of tissue	670	37.5	25.5	1087	1	P80691	Hybrid plasminogen
598	37.5	25.5	562	1	R07079	Thrombolytic prote	671	37.5	25.5	1196	1	W21015	HSV-2 strain SB5 C
599	37.5	25.5	562	1	P82580	Tissue plasminogen	672	37.5	25.5	1588	1	W35576	TFP-RI-DD ligand p
600	37.5	25.5	562	1	P82582	Tissue plasminogen	673	37.5	25.5	1602	1	W64453	Rat brain Rabsi GEP
601	37.5	25.5	562	1	P81359	Pre-pro tissue pla	674	37.5	25.2	40	1	P70205	Sequence of inhibi
602	37.5	25.5	562	1	R06655	Tissue plasminogen	675	37.5	25.2	42	1	Y12782	Human 5' EST secre
603	37.5	25.5	562	1	P80656	Tissue plasminogen	676	37.5	25.2	70	1	W76662	Elapidae modified
604	37.5	25.5	562	1	P81913	Tissue plasminogen	677	37.5	25.2	73	1	W76663	Elapidae modified
605	37.5	25.5	562	1	R06237	Novel tissue plasm	678	37.5	25.2	74	1	W76661	Elapidae modified
606	37.5	25.5	562	1	P70474	Sequence of tissue	679	37.5	25.2	95	1	W28209	Amino acid sequenc
607	37.5	25.5	562	1	R12343	T-PA with -ve char	680	37.5	25.2	207	1	R12931	Human O6-methylgua
608	37.5	25.5	562	1	R12366	T-PA with -ve char	681	37.5	25.2	217	1	W97318	A HQBQ59 polypept
609	37.5	25.5	562	1	R12423	T-PA variant havin	682	37.5	25.2	231	1	R66292	Mycobacterium bovi
610	37.5	25.5	562	1	R12424	T-PA variant havin	683	37.5	25.2	231	1	R63900	M. bovis Inha. Pol
611	37.5	25.5	562	1	R12425	T-PA variant havin	684	37.5	25.2	259	1	R24257	Alpha-ald gene pro
612	37.5	25.5	562	1	R12847	T-PA Kringle 1 dom	685	37.5	25.2	266	1	R43576	Bovine adrenocorti
613	37.5	25.5	562	1	R13441	MB103 t-PA varian	686	37.5	25.2	269	1	R66290	M. tuberculosis in
614	37.5	25.5	562	1	R12847	Sequence of human	687	37.5	25.2	269	1	R66901	M. tuberculosis in
615	37.5	25.5	562	1	P60790	Sequence of active	688	37.5	25.2	269	1	W40805	M. tuberculosis In
616	37.5	25.5	562	1	P60810	Sequence of modifi	689	37.5	25.2	269	1	W40806	M. bovis Inha prot
617	37.5	25.5	562	1	R13727	T-PA67+ mutant wit	690	37.5	25.2	296	1	R65488	Human chondromodul

691	37	25.2	296	1	R65491	Human chondromodul	764	36.5	24.8	554	1	P70436	Macrophage-colony
692	37	25.2	297	1	W37830	Bovine adrenocort	765	36.5	24.8	554	1	P81040	Sequence of human
693	37	25.2	297	1	W87867	Bovine melanocort	766	36.5	24.8	554	1	R10434	Monocyte macrophag
694	37	25.2	318	1	W00474	Soluble CD14 deriv	767	36.5	24.8	554	1	R30196	M-CSF-beta. Isola
695	37	25.2	334	1	R65486	Human chondromodul	768	36.5	24.8	554	1	R37997	Sequence of human
696	37	25.2	334	1	R65487	Human chondromodul	769	36.5	24.8	554	1	R44839	Sequence of a huma
697	37	25.2	334	1	R65489	Human chondromodul	770	36.5	24.8	554	1	R48677	Human MCSF. Transp
698	37	25.2	334	1	R65490	Human chondromodul	771	36.5	24.8	620	1	W14995	Human c-Fos induce
699	37	25.2	334	1	W44784	Human chondromodul	772	36.5	24.8	626	1	R75939	Murine myeloprolif
700	37	25.2	348	1	R98576	CD14 sequence. Pep	773	36.5	24.8	633	1	R79908	Type I MPL recepto
701	37	25.2	348	1	W00467	Soluble CD14. Modi	774	36.5	24.8	633	1	R79053	Mouse type I MPL r
702	37	25.2	348	1	W00468	Soluble CD14 deriv	775	36.5	24.8	633	1	W03513	Mouse type I MPL r
703	37	25.2	348	1	W00469	Soluble CD14 deriv	776	36.5	24.8	633	1	R98948	Mouse type I MPL r
704	37	25.2	348	1	W00470	Soluble CD14 deriv	777	36	24.5	36	1	W47064	CTLA-4 receptor cy
705	37	25.2	348	1	W00471	Soluble CD14 deriv	778	36	24.5	37	1	R39317	VI toxin. Insectic
706	37	25.2	348	1	W00472	Soluble CD14 deriv	779	36	24.5	37	1	W78452	Insecticidal pepti
707	37	25.2	348	1	W00473	Soluble CD14 deriv	780	36	24.5	37	1	W78454	Insecticidal pepti
708	37	25.2	375	1	W05316	Myelomonocytic dif	781	36	24.5	55	1	W25053	BRCA2 cancer suscep
709	37	25.2	375	1	W41693	Human CD14. Assay	782	36	24.5	166	1	W39468	Human secreted pro
710	37	25.2	375	1	W60854	Human CD14 protein	783	36	24.5	180	1	W96601	Human zinc binding
711	37	25.2	376	1	W76765	R. sphaeroides Adh	784	36	24.5	180	1	W72089	HSV-2 strain SB5 C
712	37	25.2	376	1	W72976	Rhodobacter sphaer	785	36	24.5	180	1	W72170	HSV-2 strain SB5 C
713	37	25.2	392	1	R12345	Toxoplasma gondii	786	36	24.5	183	1	W72006	Fragment of human
714	37	25.2	426	1	R05413	BUF-3 human differ	787	36	24.5	185	1	W67957	Human T cell speci
715	37	25.2	426	1	P70203	Sequence of human	788	36	24.5	187	1	R86945	Soluble CTLA4 muta
716	37	25.2	427	1	P70200	Sequence of porcine	789	36	24.5	187	1	W29728	Human CTLA4 recepto
717	37	25.2	428	1	R12352	Toxoplasma gondii	790	36	24.5	187	1	W97610	Human CTLA4 recepto
718	37	25.2	432	1	R92813	Murine interleukin	791	36	24.5	187	1	W97615	Human CTLA recepto
719	37	25.2	441	1	R99091	Murine Etl-2 gene	792	36	24.5	203	1	W97355	A partial HOQB059
720	37	25.2	467	1	W76426	Glucocorticoid ind	793	36	24.5	211	1	W87560	Human CTLA4 recept
721	37	25.2	467	1	Y05284	EGF-like homologue	794	36	24.5	212	1	R31040	Human CTLA4 recept
722	37	25.2	467	1	Y13346	Amino acid sequenc	795	36	24.5	212	1	R60134	CTLA4 receptor fus
723	37	25.2	524	1	W48893	Homo sapiens MGMT-	796	36	24.5	212	1	R7642	Full length CTLA4
724	37	25.2	543	1	W07702	Mouse ETS2 repress	797	36	24.5	212	1	R81584	Human CTLA4 recept
725	37	25.2	548	1	W07700	Human ETS2 repress	798	36	24.5	212	1	W73630	Human secreted pro
726	37	25.2	548	1	Y06927	C. albicans antige	799	36	24.5	219	1	W40086	Seq ID 91 from US5
727	37	25.2	608	1	R33646	Rabbit pre-pro ser	800	36	24.5	219	1	W59197	Seq ID 91 from US
728	37	25.2	640	1	R82249	Chloroperoxidase.	801	36	24.5	223	1	W25111	Soluble human CTLA
729	37	25.2	1061	1	R04192	Pol gene of simian	802	36	24.5	229	1	W04248	Moloney murine leu
730	37	25.2	1065	1	W33819	Arabidopsis cellul	803	36	24.5	229	1	W81136	Receptor binding r
731	37	25.2	1148	1	R07087	Renal cancer assoc	804	36	24.5	232	1	W64261	Kallikrein substr
732	37	25.2	2189	1	R05222	Antigen GX5401FL e	805	36	24.5	248	1	W89395	Modified receptor
733	37	25.2	3110	1	R71730	Merosin major subu	806	36	24.5	270	1	W15252	Asialoglycoprotein
734	37	25.2	4473	1	R97244	Virulence gene clu	807	36	24.5	287	1	W15246	Asialoglycoprotein
735	36.5	24.8	20	1	R30901	Cell adhesion poly	808	36	24.5	309	1	R64189	Murine 4-1BB-L pol
736	36.5	24.8	79	1	W28150	Staphylococcus aur	809	36	24.5	309	1	W14054	Human MAT1. Novel
737	36.5	24.8	121	1	Y07785	Human secreted pro	810	36	24.5	309	1	W26656	Murine 4-1BB ligan
738	36.5	24.8	177	1	R81429	Hepatitis GB virus	811	36	24.5	312	1	W09778	Human MAT1. Assemb
739	36.5	24.8	232	1	R29716	Guinea pig MBP-1.	812	36	24.5	339	1	R42841	Human haematopoiet
740	36.5	24.8	233	1	W40075	Guinea pig eosinop	813	36	24.5	341	1	W73330	Minos-3 protein se
741	36.5	24.8	286	1	Y05220	Kringiel protein s	814	36	24.5	343	1	W80577	Partial human sequ
742	36.5	24.8	299	1	R04909	Rat binding protei	815	36	24.5	362	1	W97364	Amino acid sequenc
743	36.5	24.8	400	1	R98992	Human M-CSF derive	816	36	24.5	365	1	R20103	ERK1. DNA encoding
744	36.5	24.8	408	1	W10082	N-delta-3/C-delta-	817	36	24.5	365	1	W80571	Human heart O-fuco
745	36.5	24.8	429	1	Y06993	Rat eukaryotic tra	818	36	24.5	382	1	W05322	Human mucosal addr
746	36.5	24.8	482	1	R75941	Soluble murine MPL	819	36	24.5	382	1	Y02069	Human MADCAM-1 pro
747	36.5	24.8	519	1	W10072	N-delta 3 LCSF. DN	820	36	24.5	393	1	R99416	Aminopeptidase pre
748	36.5	24.8	519	1	W10084	N-delta-3/C-delta-	821	36	24.5	397	1	W80573	Human O-fucosyltra
749	36.5	24.8	519	1	W10085	N-delta-3/Lys52Glu	822	36	24.5	403	1	W81588	Protein encoded by
750	36.5	24.8	520	1	R12111	Human M-CSF deriva	823	36	24.5	404	1	R78621	Chicken GalNAc-4alp
751	36.5	24.8	536	1	P95032	Long form of colon	824	36	24.5	418	1	R90287	Pigment epithelium
752	36.5	24.8	536	1	P90567	Colony stimulating	825	36	24.5	455	1	W39171	Human PKD1 protein
753	36.5	24.8	536	1	R05487	Long form of human	826	36	24.5	461	1	R92220	Calcitonin gene-re
754	36.5	24.8	536	1	P80764	Deduced amino acid	827	36	24.5	469	1	R55340	Sequence of envelo
755	36.5	24.8	536	1	W10066	Human long form CS	828	36	24.5	469	1	W26321	MoMLV ecotropic gp
756	36.5	24.8	536	1	W22613	Human colony stimu	829	36	24.5	471	1	Y05372	Human HCMV inducib
757	36.5	24.8	536	1	W35750	Human colony stimu	830	36	24.5	472	1	R92753	Human calcitonin r
758	36.5	24.8	536	1	W36141	Human long form of	831	36	24.5	533	1	W56484	zea mays waxy gene
759	36.5	24.8	536	1	W93901	Human pCCSF-4 prot	832	36	24.5	534	1	W25031	Partial BRCA2 can
760	36.5	24.8	551	1	W27150	HMG-CoA reductase	833	36	24.5	596	1	W76820	Human Tcf-4E prote
761	36.5	24.8	554	1	P91044	Human macrophage c	834	36	24.5	624	1	W85015	Erk1-green floures
762	36.5	24.8	554	1	P91898	Sequence of a macr	835	36	24.5	648	1	R22561	Mouse mutant C-raf
763	36.5	24.8	554	1	P90123	Recombinant human	836	36	24.5	648	1	R22563	Mouse mutant C-raf

837	36	24.5	648	1	R22559	Mouse c-raf-1. Det	104	35.5	24.1	189	1	P50328	Bovine interferon
838	36	24.5	648	1	R25277	Human c-raf-1. Det	104	35.5	24.1	189	1	P30073	Sequence of bovine
839	36	24.5	648	1	R22560	Mouse mutant c-raf	104	35.5	24.1	189	1	W83895	Bovine interferon-
840	36	24.5	648	1	R22562	Mouse mutant c-raf	104	35.5	24.1	189	1	W73224	Bovine interferon-
841	36	24.5	648	1	R98215	Human Raf1 Kinase.	104	35.5	24.1	264	1	W40088	Seq ID 93 from US
842	36	24.5	648	1	W13107	Human Raf-1. Compl	104	35.5	24.1	264	1	W59199	Seq ID 93 from US
843	36	24.5	648	1	W17046	Mutant mouse c-raf	104	35.5	24.1	288	1	R81499	BRCA1 mutant from
844	36	24.5	648	1	W17047	Mutant mouse c-raf	104	35.5	24.1	288	1	R99442	BRCA1 allele #388
845	36	24.5	648	1	W17048	Mutant mouse c-raf	104	35.5	24.1	312	1	W10005	Protein encoded by
846	36	24.5	648	1	W17049	Mutant mouse c-raf	104	35.5	24.1	312	1	W10005	Protein encoded by
847	36	24.5	648	1	W17045	Mouse c-raf 1 prot	104	35.5	24.1	312	1	W79891	Tumorigenic BRCA1
848	36	24.5	648	1	W17044	Mouse c-raf 1 prot	104	35.5	24.1	344	1	R81501	BRCA1 mutant from
849	36	24.5	648	1	W17044	Human c-raf 1 prot	104	35.5	24.1	395	1	R81502	BRCA1 mutant from
850	36	24.5	648	1	W62220	Raf-1 protein. Det	104	35.5	24.1	469	1	R04944	Equine neraminidase
851	36	24.5	648	1	W30668	Mutant mouse c-raf	104	35.5	24.1	470	1	R70712	Equine influenza v
852	36	24.5	648	1	W30669	Mutant mouse c-raf	104	35.5	24.1	478	1	R81503	BRCA1 mutant from
853	36	24.5	648	1	W30664	Human c-raf-1 prot	104	35.5	24.1	479	1	R81036	Modified tissue pl
854	36	24.5	648	1	W30665	Mouse c-raf-1 prot	104	35.5	24.1	481	1	R81492	BRCA1 mutant from
855	36	24.5	648	1	W30666	Mutant mouse c-raf	104	35.5	24.1	527	1	R09266	t-PA variant R267A
856	36	24.5	648	1	W30667	Mutant mouse c-raf	104	35.5	24.1	527	1	R70891	Human t-PA variant
857	36	24.5	648	1	W56111	Homo sapiens GST-R	104	35.5	24.1	527	1	R70856	Human t-PA variant
858	36	24.5	648	1	W57320	Staphylococcus aur	104	35.5	24.1	559	1	R13157	T-PA variant havin
859	36	24.5	711	1	W57321	Topoisomerase I ex	104	35.5	24.1	562	1	R12426	T-PA variant havin
860	36	24.5	712	1	W44893	Chimeric polyproli	104	35.5	24.1	671	1	R81504	BRCA1 mutant from
861	36	24.5	752	1	W44894	Human transcrip	104	35.5	24.1	685	1	W31274	Mouse frizzled-8 p
862	36	24.5	757	1	W44894	Polyproline-contai	104	35.5	24.1	730	1	R80618	Human Bone Morphog
863	36	24.5	772	1	R23873	Human alpha 5 (IV)	104	35.5	24.1	730	1	W13669	C-proteinase encod
864	36	24.5	773	1	W44887	Chimeric polyproli	104	35.5	24.1	734	1	R81506	BRCA1 mutant from
865	36	24.5	783	1	W44895	Polyproline-contai	104	35.5	24.1	765	1	R99443	BRCA1 allele #6401
866	36	24.5	786	1	R47066	Sequence of Crypto	104	35.5	24.1	765	1	W10006	Protein encoded by
867	36	24.5	788	1	W44888	Chimeric polyproli	104	35.5	24.1	765	1	W79892	Tumorigenic BRCA1
868	36	24.5	816	1	W68094	Mouse neuronal PAS	104	35.5	24.1	788	1	W75919	C-proteinase seque
869	36	24.5	829	1	W44889	Chimeric polyproli	104	35.5	24.1	793	1	W40117	Human GABA-BR1a/b
870	36	24.5	839	1	R22275	Bovine liver GPI-P	104	35.5	24.1	798	1	R81491	BRCA1 mutant from
871	36	24.5	842	1	W56698	Neuronal migration	104	35.5	24.1				
872	36	24.5	880	1	W44884	Chimeric polyproli	104	35.5	24.1				
873	36	24.5	882	1	W44899	Chimeric polyproli	104	35.5	24.1				
874	36	24.5	882	1	W44901	Chimeric polyproli	104	35.5	24.1				
875	36	24.5	892	1	W44903	Chimeric polyproli	104	35.5	24.1				
876	36	24.5	899	1	W44886	Chimeric polyproli	104	35.5	24.1				
877	36	24.5	916	1	W85612	Secreted protein c	104	35.5	24.1				
878	36	24.5	919	1	W18580	Potato alpha-gluco	104	35.5	24.1				
879	36	24.5	926	1	W44883	Chimeric polyproli	104	35.5	24.1				
880	36	24.5	930	1	W44885	Chimeric polyproli	104	35.5	24.1				
881	36	24.5	979	1	R34783	Human CENP-C anti	104	35.5	24.1				
882	36	24.5	997	1	R15345	Mouse bullous pem	104	35.5	24.1				
883	36	24.5	1132	1	R77417	Human cell cycle p	104	35.5	24.1				
884	36	24.5	1142	1	W81546	Tumour rejection a	104	35.5	24.1				
885	36	24.5	1203	1	R41568	Human endothelial	104	35.5	24.1				
886	36	24.5	1203	1	W57315	Human endothelial	104	35.5	24.1				
887	36	24.5	1237	1	R13791	E75A protein. DNA	104	35.5	24.1				
888	36	24.5	1376	1	R93601	Kaposi's sarcoma a	104	35.5	24.1				
889	36	24.5	1376	1	R97835	Kaposi's sarcoma a	104	35.5	24.1				
890	36	24.5	1464	1	W42632	Human transcrip	104	35.5	24.1				
891	36	24.5	1607	1	W50897	Mouse laminin Gl c	104	35.5	24.1				
892	36	24.5	1609	1	W50898	Human laminin Gl c	104	35.5	24.1				
893	36	24.5	2321	1	W49698	Human Notch3 prote	104	35.5	24.1				
894	36	24.5	2329	1	W25038	Partial BRCA2 canc	104	35.5	24.1				
895	36	24.5	3418	1	W23287	Human breast and o	104	35.5	24.1				
896	36	24.5	3418	1	W19211	Human breast canc	104	35.5	24.1				
897	36	24.5	3418	1	Y04355	Human BRCA2 (om12)	104	35.5	24.1				
898	36	24.5	3418	1	Y04356	Human BRCA2 (om13)	104	35.5	24.1				
899	36	24.5	3418	1	Y04357	Human BRCA2 (om14)	104	35.5	24.1				
900	36	24.5	3418	1	Y04358	Human BRCA2 (om15)	104	35.5	24.1				
901	36	24.5	3418	1	Y04354	Human BRCA2 (om11)	104	35.5	24.1				
902	36	24.5	4302	1	W00870	Polycystic kidney	104	35.5	24.1				
903	36	24.5	4302	1	W23830	Human PKD1 protein	104	35.5	24.1				
904	36	24.5	4302	1	W33396	Human PKD1 polyep	104	35.5	24.1				
905	36	24.5	4339	1	R87539	Polycystic kidney	104	35.5	24.1				
906	36	24.5	4472	1	W22601	Tyrosine synthase	104	35.5	24.1				
907	35.5	24.1	51	1	R49335	Camel Ig 2-heavy c	104	35.5	24.1				
908	35.5	24.1	33	1	R49721	Sequence of a pept	104	35.5	24.1				
909	35.5	24.1	65	1	R39669	C-terminal Kunitz-	104	35.5	24.1				

ALIGNMENTS

RESULT 1

W99544
ID W99544 standard; Protein; 116 AA.
AC W99544;
DT 08-JUN-1999 (first entry)
DE GPH analogue alpha-subunit alpha-V4C.
KW Analogue; heterodimeric; glycoprotein hormone; hCG; hLH; hFSH; hTSH;
KW human chorionic gonadotropin; human luteinising hormone; disulphide bond;
KW human follicle stimulating hormone; human thyroid stimulating hormone;
KW stability; primer; amplification; PCR; mutation.
OS Homo sapiens.
OS Synthetic.
FN W09858957-A2.
PD 30-DEC-1998.
PF 25-JUN-1998; U13070.
PR 25-JUN-1997; U13070.
PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PA (MCIN-) MCINIS P G.
PI Moyle WR;
DR WPI; 99-081219/07.
PT New stabilised glycoprotein hormones - particularly hCG, hLH, hFSH
PT or hTSH, have an intersubunit disulphide crosslink between the
PT alpha- and beta-subunits to improve stability
PS Disclosure; Fig 34A; 139pp; English.
CC The invention relates to the production of analogues of a heterodimeric
CC subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
CC (hCG), human luteinising hormone (hLH), human follicle stimulating
CC hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
CC muteins, which are modified to contain an intersubunit disulphide bond,
CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
CC improved stability, the analogue retaining at least a portion of the
CC bioactivity for the corresponding native GPH receptor. This sequence
CC represents a mutant GPH-alpha subunit used for the generation of the
CC modified GPHs. The improved analogues are designed specifically to
CC reduce perturbation of the 3-dimensional structure of the hormone.

CC subunit glycoprotein hormone (GPH) e.g. human chorionic gonadotropin
 CC (hCG), human luteinising hormone (hLH), human follicle stimulating
 CC hormone (hFSH), human thyroid stimulating hormone (hTSH), and functional
 CC muteins, which are modified to contain an intersubunit disulphide bond,
 CC between an alpha-subunit cysteine and a beta-subunit cysteine, for
 CC improved stability, the analogue retaining at least a portion of the
 CC bioactivity for the corresponding native GPH receptor. This sequence
 CC represents a mutant hCG-alpha subunit used for the generation of the
 CC modified GPHs. The improved analogues are designed specifically to
 CC reduce perturbation of the 3-dimensional structure of the hormone,
 CC thereby creating greater likelihood that the dimer will be formed in vivo
 CC and the formed dimer will have affinity for the native receptors and have
 CC agonistic activity on them. The changes stabilise the GPHs and prolong
 CC the biological activities of the hormones. The analogues can have uses
 CC as for the native GPHs.
 SQ Sequence 116 AA;

Query Match 93.2%; Score 137; DB 1; Length 116;
 Best Local Similarity 96.2%; Pred. No. 2.7e-12;
 Matches 25; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 apdvdcpectlqenpfssqgapil 26
 ||||| ||||| ||||| ||||| |||||
 Db 25 APDVQDSPECTLQENPFSSQPGAPIL 50

RESULT 5

ID R15197 standard; Protein; 68 AA.
 AC R15197;
 DT 11-FEB-1992 (first entry)
 DE hCG/bCG alpha subunit chimera, H19.
 KW Glycoprotein hormone; human chorionic gonadotropin; bovine.
 OS Homo sapiens.
 OS Bos taurus.
 PN W09116922-A.
 PD 14-NOV-1991.
 PF 07-MAY-1991; US-520703.
 PR 08-MAY-1990; US-520703.
 PA (UYNE-) UNIV MED NEW JERSEY.
 PI Campbell RK, Moyle WR;
 DR WPI; 91-353528/48.
 DT New glyco-protein hormone analogues - for inducing fertility as
 PT immuno-castration agents, for suppressing reproductive system
 PT development and as immuno-contragestive vaccines.
 PS Table IX; Page 68; 94pp; English.
 CC The sequence is an analogue of mature hCG alpha subunit having
 CC several residues replaced by the corresponding residues in the
 CC bovine sequence and residues 34-61 deleted. The subunit will
 CC combine with hCG beta subunit and bind LH receptors. The chimera
 CC may be useful for the construction of hormone dimers that can be
 CC used as immunogens such that antibodies cross-reactive with other
 CC naturally occurring glycoprotein hormones will not be produced. It
 CC may also provide an alternative sequence to place in human hormones
 CC to alter the properties of the molecule.
 CC See R15043, R15061-R15125 and R15161-R15198.
 SQ Sequence 68 AA;

Query Match 82.3%; Score 121; DB 1; Length 68;
 Best Local Similarity 79.3%; Pred. No. 2.6e-10;
 Matches 23; Conservative 1; Mismatches 1; Indels 4; Gaps 1;

Qy 2 pd----vqdcpectlqenpfssqgapil 26
 || : ||||| ||||| ||||| |||||
 Db 2 PDGEFTMQGPECTLQENPFSSQPGAPIL 30

RESULT 6

R15182
 ID R15182 standard; Protein; 96 AA.
 AC R15182;

DT 11-FEB-1992 (first entry)
 DE hCG/bCG alpha subunit chimera, H4.
 KW Glycoprotein hormone; human chorionic gonadotropin; bovine.
 OS Homo sapiens.
 OS Bos taurus.
 PN W09116922-A.
 PD 14-NOV-1991.
 PF 07-MAY-1991; U03162.
 PR 08-MAY-1990; US-520703.
 PA (UYNE-) UNIV MED NEW JERSEY.
 PI Campbell RK, Moyle WR;
 DR WPI; 91-353528/48.
 DT New glyco-protein hormone analogues - for inducing fertility as
 PT immuno-castration agents, for suppressing reproductive system
 PT development and as immuno-contragestive vaccines.
 PS Table IX; Page 68; 94pp; English.
 CC The sequence is an analogue of mature hCG alpha subunit having
 CC several residues replaced by the corresponding residues in the
 CC bovine sequence. The subunit will combine with hCG beta subunit
 CC and bind LH receptors. The chimera may be useful for the
 CC construction of hormone dimers that can be used as immunogens such
 CC that antibodies cross-reactive with other naturally occurring glyco-
 CC protein hormones will not be produced. It may also provide an
 CC alternative sequence to place in human hormones to alter the proper-
 CC ties of the molecule.
 CC See R15043, R15061-R15125 and R15161-R15198.
 SQ Sequence 96 AA;

Query Match 82.3%; Score 121; DB 1; Length 96;
 Best Local Similarity 79.3%; Pred. No. 3.7e-10;
 Matches 23; Conservative 1; Mismatches 1; Indels 4; Gaps 1;

Qy 2 pd----vqdcpectlqenpfssqgapil 26
 || : ||||| ||||| ||||| |||||
 Db 2 PDGEFTMQGPECTLQENPFSSQPGAPIL 30

RESULT 7

R15183
 ID R15183 standard; Protein; 96 AA.
 AC R15183;
 DT 11-FEB-1992 (first entry)
 DE hCG/bCG alpha subunit chimera, H5.
 KW Glycoprotein hormone; human chorionic gonadotropin; bovine.
 OS Homo sapiens.
 OS Bos taurus.
 PN W09116922-A.
 PD 14-NOV-1991.
 PF 07-MAY-1991; U03162.
 PR 08-MAY-1990; US-520703.
 PA (UYNE-) UNIV MED NEW JERSEY.
 PI Campbell RK, Moyle WR;
 DR WPI; 91-353528/48.
 DT New glyco-protein hormone analogues - for inducing fertility as
 PT immuno-castration agents, for suppressing reproductive system
 PT development and as immuno-contragestive vaccines.
 PS Table IX; Page 68; 94pp; English.
 CC The sequence is an analogue of mature hCG alpha subunit having
 CC several residues at the N-terminal replaced by the corresponding
 CC residues in the bovine sequence. The subunit will combine with hCG
 CC beta subunit and bind LH receptors. The chimera may be useful for
 CC the construction of hormone dimers that can be used as immunogens such
 CC that antibodies cross-reactive with other naturally occurring glyco-
 CC protein hormones will not be produced. It may also provide an
 CC alternative sequence to place in human hormones to alter the proper-
 CC ties of the molecule.
 CC See R15043, R15061-R15125 and R15161-R15198.
 SQ Sequence 96 AA;

Query Match 82.3%; Score 121; DB 1; Length 96;
 Best Local Similarity 79.3%; Pred. No. 3.7e-10;

Matches 23; Conservative 1; Mismatches 1; Indels 4; Gaps 1;

QY 2 pd----vqdcpectqlenpfssqgapil 26
|| : ||||| ||||| ||||| |||||
Db 2 PDGEFTMQGCECTQLQENPFSSQPGAPIL 30

RESULT 8

R15191
ID R15191 standard; Protein; 96 AA.
AC R15191;
DT 11-FEB-1992 (first entry)
DE hCG/bCG alpha subunit chimera, H13.
KW Glycoprotein hormone; human chorionic gonadotropin; bovine.
OS Homo sapiens.
PN WO9116922-A.
PD 14-NOV-1991.
PF 07-MAY-1991; U03162.
PR 08-MAY-1990; US-520703.
PA (UYNE-) UNIV MED NEW JERSEY.
PI Campbell RK, Moyle WR;
DR WPI; 91-353528/48.
PT New glyco-protein hormone analogues - for inducing fertility as
PT immuno-castration agents, for suppressing reproductive system
PT development and as immuno-contraceptive vaccines.
PS Table IX; Page 68; 94pp; English.
CC The sequence is an analogue of mature hCG alpha subunit having
CC several residues at the N-terminal replaced by the corresponding
CC residues in the bovine sequence. The subunit will combine with hCG
CC beta subunit and bind LH receptors. The chimera may be useful for
CC the construction of hormone dimers that can be used as immunogens such
CC that antibodies cross-reactive with other naturally occurring glyco-
CC protein hormones will not be produced. It may also provide an
CC alternative sequence to place in human hormones to alter the proper-
CC ties of the molecule.
CC See R15043, R15061-R15125 and R15161-R15198.
SQ Sequence 96 AA;

Query Match 82.3%; Score 121; DB 1; Length 96;

Best Local Similarity 79.3%; Pred. No. 3.7e-10;
Matches 23; Conservative 1; Mismatches 1; Indels 4; Gaps 1;

QY 2 pd----vqdcpectqlenpfssqgapil 26
|| : ||||| ||||| ||||| |||||
Db 2 PDGEFTMQGCECTQLQENPFSSQPGAPIL 30

RESULT 9

W33357
ID W33357 standard; Protein; 256 AA.
AC W33357;
DT 19-MAR-1998 (first entry)
DE TBP(20-161)/hCG-alpha fusion protein.
KW Fusion protein; thrombopoietin; TPO; human chorionic gonadotropin;
KW alpha subunit; hCG-alpha.
OS Homo sapiens.
PN WO9730161-A1.
PD 21-AUG-1997.
PF 20-FEB-1997; U02315.
PR 20-FEB-1996; US-011936.
PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PI Campbell RK, Chappel SC, Jameson BA;
DR WPI; 97-425036/39.
DR N-PSDB; T94007.
PT Hybrid dimeric protein comprising two co-expressed units - each
PT based on receptor or ligand and a subunit of a heterodimeric
PT hormone, especially FSH, for inducing follicular maturation
PS Example; Pages 32-33; 60pp; English.
CC A novel fusion protein comprises 2 dimer forming co-expressed amino
CC acid sequences, each consisting of a homodimeric or heterodimeric
CC receptor chain or ligand, with ligand-receptor binding activity,

CC bound directly or via a peptide linker to a subunit of a
CC heterodimeric protein hormone capable of forming a heterodimer with
CC the hormone's other subunits. The fusion protein, e.g. the
CC thrombopoietin (TPO)/human chorionic gonadotropin-alpha subunit
CC (hCG-alpha) fusion protein denoted by the present sequence,
CC significantly increases the biological activity of the hormone
CC component, reducing the requirement for hormone itself and the
CC number of injections needed.
SQ Sequence 256 AA;

Query Match 78.2%; Score 115; DB 1; Length 256;

Best Local Similarity 100.0%; Pred. No. 7.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 cpectqlenpfssqgapil 26
||||| ||||| ||||| |||||
Db 171 CPECTQLQENPFSSQPGAPIL 190

RESULT 10

W33359
ID W33359 standard; Protein; 285 AA.
AC W33359;
DT 19-MAR-1998 (first entry)
DE TBP(20-190)/hCG-alpha fusion protein.
KW Fusion protein; thrombopoietin; TPO; human chorionic gonadotropin;
KW alpha subunit; hCG-alpha.
OS Homo sapiens.
PN WO9730161-A1.
PD 21-AUG-1997; U02315.
PF 20-FEB-1997; U02315.
PR 20-FEB-1996; US-011936.
PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.
PI Campbell RK, Chappel SC, Jameson BA;
DR WPI; 97-425036/39.
DR N-PSDB; T94021.
PT Hybrid dimeric protein comprising two co-expressed units - each
PT based on receptor or ligand and a subunit of a heterodimeric
PT hormone, especially FSH, for inducing follicular maturation
PS Example; Pages 37-38; 60pp; English.
CC A novel fusion protein comprises 2 dimer forming co-expressed amino
CC acid sequences, each consisting of a homodimeric or heterodimeric
CC receptor chain or ligand, with ligand-receptor binding activity,
CC bound directly or via a peptide linker to a subunit of a
CC heterodimeric protein hormone capable of forming a heterodimer with
CC the hormone's other subunits. The fusion protein, e.g. the
CC thrombopoietin (TPO)/human chorionic gonadotropin-alpha subunit
CC (hCG-alpha) fusion protein denoted by the present sequence,
CC significantly increases the biological activity of the hormone
CC component, reducing the requirement for hormone itself and the
CC number of injections needed.
SQ Sequence 285 AA;

Query Match 78.2%; Score 115; DB 1; Length 285;

Best Local Similarity 100.0%; Pred. No. 8.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 cpectqlenpfssqgapil 26
||||| ||||| ||||| |||||
Db 200 CPECTQLQENPFSSQPGAPIL 219

RESULT 11

R15180
ID R15180 standard; Protein; 92 AA.
AC R15180;
DT 11-FEB-1992 (first entry)
DE hCG/bCG alpha subunit chimera, H2.
KW Glycoprotein hormone; human chorionic gonadotropin; bovine.
OS Homo sapiens.
OS Bos taurus.

CC See R15043, R15061-R15125 and R15161-R15198.
SQ Sequence 96 AA;

Query Match 72.1%; Score 106; DB 1; Length 96;
Best Local Similarity 71.4%; Pred. No. 4.7e-08;
Matches 20; Conservative 2; Mismatches 2; Indels 4; Gaps 1;

Qy 2 pd-----vqdcpctqlqenpfssqgapi 25
|| : ||||| ||||| ||||| : |||||
Db 2 PDGFTMQGCEPTLQENPFSSKPDAP1 29

RESULT 15

R15190
ID R15190 standard; Protein; 96 AA.

AC R15190;
DT 11-FEB-1992 (first entry)
DE hCG/bCG alpha subunit chimera, H12;
KW Glycoprotein hormone; human chorionic gonadotropin; bovine.

OS Homo sapiens.

OS Bos taurus.

PN W09116922-A.

PD 14-NOV-1991.

PF 07-MAY-1991; U03162.

PR 08-MAY-1990; US-520703.

PA (UYNE-) UNIV MED NEW JERSEY.

PI Campbell RK, Moyie WR;

DR WPI; 91-353528/48.

PT New glyco-protein hormone analogues - for inducing fertility as

PT immuno-castration agents, for suppressing reproductive system

PT development and as immuno-contragestive vaccines.

PS Table IX; Page 68; 94pp; English.

CC The sequence is an analogue of mature hCG alpha subunit having
CC several residues at the N-terminal replaced by the corresponding
CC residues in the bovine sequence. The subunit will combine with hCG
CC beta subunit and bind LH receptors. The chimera may be useful for

CC the construction of hormone dimers that can be used as immunogens such

CC that antibodies cross-reactive with other naturally occurring glyco-

CC protein hormones will not be produced. It may also provide an

CC alternative sequence to place in human hormones to alter the proper-

CC ties of the molecule.

CC See R15043, R15061-R15125 and R15161-R15198.

CC Sequence 96 AA;

Query Match 72.1%; Score 106; DB 1; Length 96;
Best Local Similarity 71.4%; Pred. No. 4.7e-08;
Matches 20; Conservative 2; Mismatches 2; Indels 4; Gaps 1;

Qy 2 pd-----vqdcpctqlqenpfssqgapi 25
|| : ||||| ||||| ||||| : |||||
Db 2 PDGFTMQGCEPTLQENPFSSKPDAP1 29

RESULT 16

R15188
ID R15188 standard; Protein; 96 AA.

AC R15188;
DT 11-FEB-1992 (first entry)
DE hCG/bCG alpha subunit chimera, H10.
KW Glycoprotein hormone; human chorionic gonadotropin; bovine.

OS Homo sapiens.

OS Bos taurus.

PN W09116922-A.

PD 14-NOV-1991.

PF 07-MAY-1991; U03162.

PR 08-MAY-1990; US-520703.

PA (UYNE-) UNIV MED NEW JERSEY.

PI Campbell RK, Moyie WR;

DR WPI; 91-353528/48.

PT New glyco-protein hormone analogues - for inducing fertility as

PT immuno-castration agents, for suppressing reproductive system

PT development and as immuno-contragestive vaccines.
PS Table IX; Page 68; 94pp; English.

CC The sequence is an analogue of mature hCG alpha subunit having
CC several residues at the N-terminal replaced by the corresponding
CC residues in the bovine sequence. The subunit will combine with hCG
CC beta subunit and bind LH receptors. The chimera may be useful for

CC the construction of hormone dimers that can be used as immunogens such

CC that antibodies cross-reactive with other naturally occurring glyco-

CC protein hormones will not be produced. It may also provide an

CC alternative sequence to place in human hormones to alter the proper-

CC ties of the molecule.

CC See R15043, R15061-R15125 and R15161-R15198.

CC Sequence 96 AA;

Query Match 71.4%; Score 105; DB 1; Length 96;
Best Local Similarity 71.4%; Pred. No. 6.4e-08;
Matches 20; Conservative 2; Mismatches 2; Indels 4; Gaps 1;

Qy 2 pd-----vqdcpctqlqenpfssqgapi 25
|| : ||||| ||||| ||||| : |||||
Db 2 PDGFTMQGCEPTLQENPFSSKPRAP1 29

RESULT 17

R15187
ID R15187 standard; Protein; 96 AA.

AC R15187;

DT 11-FEB-1992 (first entry)

DE hCG/bCG alpha subunit chimera, H9.

KW Glycoprotein hormone; human chorionic gonadotropin; bovine.

OS Homo sapiens.

OS Bos taurus.

PN W09116922-A.

PD 14-NOV-1991.

PF 07-MAY-1991; U03162.

PR 08-MAY-1990; US-520703.

PA (UYNE-) UNIV MED NEW JERSEY.

PI Campbell RK, Moyie WR;

DR WPI; 91-353528/48.

PT New glyco-protein hormone analogues - for inducing fertility as

PT immuno-castration agents, for suppressing reproductive system

PT development and as immuno-contragestive vaccines.

PS Table IX; Page 68; 94pp; English.

CC The sequence is an analogue of mature hCG alpha subunit having
CC several residues at the N-terminal replaced by the corresponding
CC residues in the bovine sequence. The subunit will combine with hCG
CC beta subunit and bind LH receptors. The chimera may be useful for

CC the construction of hormone dimers that can be used as immunogens such

CC that antibodies cross-reactive with other naturally occurring glyco-

CC protein hormones will not be produced. It may also provide an

CC alternative sequence to place in human hormones to alter the proper-

CC ties of the molecule.

CC See R15043, R15061-R15125 and R15161-R15198.

CC Sequence 96 AA;

Query Match 68.0%; Score 100; DB 1; Length 96;
Best Local Similarity 65.5%; Pred. No. 3.2e-07;
Matches 19; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

Qy 2 pd-----vqdcpctqlqenpfssqgapi 26
|| : ||||| ||||| ||||| : |||||
Db 2 PDGFTMQGCEPTLQENPFSSQPGAPIL 30

RESULT 18

P90961

ID P90961 standard; peptide; 15 AA.

AC P90961;

DT 07-FEB-1990 (first entry)

DE Synthetic glycoprotein hormone alpha subunit peptide.

KW Glycoprotein hormone; alpha subunit; thyroid stimulating immunoglobulin;

KW Graves disease.
 PN WO8908663-A.
 PD 21-SEP-1989.
 PF 13-MAR-1988; U01007.
 PR 17-MAR-1988; US-169375.
 PA (MAYO-) Mayo Found. Med. Educ.
 PI Ryan RJ, McCormick DJ, Charlesworth MC, Morris JC;
 DR WPI: 89-292494/40.
 PT Polypeptide corresp. to glycoprotein hormone alpha-subunit - useful for
 PS Disclosure; page 7; 40pp; English.
 CC Inhibiting thyroid stimulating immunoglobulin in Graves disease.
 CC The peptide corresp. to amino acids 1 to 15 of the human glycoprotein
 CC hormone subunit, and is one of a series of 10 synthetic peptides each
 CC overlapping each other by 5 aa's together covering the complete 92 aa
 CC sequence of the alpha subunit. The peptides are used to inhibit TSH,
 CC block TSH, LH and hCG binding to receptors, and inhibit TSH-mediated cAMP
 CC generation.
 CC See also P90962-P90970.
 SQ Sequence 15 AA;

Query Match 59.2%; Score 87; DB 1; Length 15;
 Best Local Similarity 100.0%; Pred. No. 3e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 apdvqdcptqlqen 15
 |||||
 DB 1 APDVQDCPTQLQEN 15

RESULT 19

R15184
 ID R15184 standard; Protein; 96 AA.
 AC R15184;
 DT 11-FEB-1992 (first entry)
 DE hCG/bCG alpha subunit chimera, H6.
 KW Glycoprotein hormone; human chorionic gonadotropin; bovine.
 OS Homo sapiens.
 OS Bos taurus.
 PN WO89116922-A.
 PD 14-NOV-1991.
 PF 07-MAY-1991; U03162.
 PR 08-MAY-1990; US-520703.
 PA (UYNE-) UNIV MED NEW JERSEY.
 PI Campbell RK, Moyle WR;
 DR WPI: 91-353528/48.
 PT New glyco-protein hormone analogues - for inducing fertility as
 PT immuno-castration agents, for suppressing reproductive system
 PT development and as immuno-contragestive vaccines.
 PS Table IX; Page 68; 94pp; English.
 CC The sequence is an analogue of mature hCG alpha subunit having
 CC several residues replaced by the corresponding residues in the
 CC bovine sequence. The subunit will combine with hCG beta subunit
 CC and bind LH receptors. The chimera may be useful for the
 CC construction of hormone dimers that can be used as immunogens such
 CC that antibodies cross-reactive with other naturally occurring glyco-
 CC protein hormones will not be produced. It may also provide an
 CC alternative sequence to place in human hormones to alter the proper-
 CC ties of the molecule.
 CC See R15043, R15061-R15125 and R15161-R15198.
 SQ Sequence 96 AA;

Query Match 57.8%; Score 85; DB 1; Length 96;
 Best Local Similarity 57.1%; Pred. No. 4e-05;
 Matches 16; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

QY 2 pd----vqdcptqlqenpfsspgapi 25
 || : |||| : || : || : ||
 DB 2 PDGEFTMQGCECKLENKYFSKPDAP1 29

RESULT 20

R15186
 ID R15186 standard; Protein; 96 AA.
 AC R15186;
 DT 11-FEB-1992 (first entry)
 DE hCG/bCG alpha subunit chimera, H8.
 KW Glycoprotein hormone; human chorionic gonadotropin; bovine.
 OS Homo sapiens.
 OS Bos taurus.
 PN WO89116922-A.
 PD 14-NOV-1991.
 PF 07-MAY-1991; U03162.
 PR 08-MAY-1990; US-520703.
 PA (UYNE-) UNIV MED NEW JERSEY.
 PI Campbell RK, Moyle WR;
 DR WPI: 91-353528/48.
 PT New glyco-protein hormone analogues - for inducing fertility as
 PT immuno-castration agents, for suppressing reproductive system
 PT development and as immuno-contragestive vaccines.
 PS Table IX; Page 68; 94pp; English.
 CC The sequence is an analogue of mature hCG alpha subunit having
 CC several residues at the N-terminal replaced by the corresponding
 CC residues in the bovine sequence. The subunit will combine with hCG
 CC beta subunit and bind LH receptors. The chimera may be useful for
 CC the construction of hormone dimers that can be used as immunogens such
 CC that antibodies cross-reactive with other naturally occurring glyco-
 CC protein hormones will not be produced. It may also provide an
 CC alternative sequence to place in human hormones to alter the proper-
 CC ties of the molecule.
 CC See R15043, R15061-R15125 and R15161-R15198.
 SQ Sequence 96 AA;

Query Match 57.8%; Score 85; DB 1; Length 96;
 Best Local Similarity 57.1%; Pred. No. 4e-05;
 Matches 16; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

QY 2 pd----vqdcptqlqenpfsspgapi 25
 || : |||| : || : || : ||
 DB 2 PDGEFTMQGCECKLENKYFSKPDAP1 29

RESULT 21

R10039
 ID R10039 standard; Protein; 120 AA.
 AC R10039;
 DT 10-APR-1991 (first entry)
 DE Ovine FSH alpha subunit precursor encoded by clone 3.2.
 KW Follicle stimulating hormone; ovulation.
 OS Ovis ammon aries.
 FH Key Location/Qualifiers
 FT peptide 1..24
 FT /label= signal sequence
 FT protein 24..120
 FT /label= FSH alpha sub-unit
 PN EP-404458-A.
 PD 27-DEC-1990.
 PF 15-JUN-1990; 306561.
 PR 19-JUN-1989; AU-004799.
 PA (BUNG-) BUNGE AUSTR PTY LTD.
 PI Adams TE, Brandon MR;
 DR WPI: 91-001487/01.
 DR N-PSDB; Q10059.
 PT Recombinant ovine follicle stimulating hormone - used for inducing or
 PT increasing ovulation in female animals and improving yield of ova.
 PS Disclosure; Fig 1b; 22pp; English.
 CC The clone was isolated from a cDNA library prepd. from RNA ex-
 CC tracted from lamb pituitary glands. The DNA is used to produce
 CC the protein by recombinant techniques. The protein can be used
 CC control ovulation and improve embryo viability in sheep and goats.
 CC See also R10038.
 SQ Sequence 120 AA;

PN J10036398-A.
PD 10-FEB-1998.
PF 25-JUL-1996; 196009.
PR 25-JUL-1996; JP-196009.
PA (ELED) DENKI KAGAKU KOGYO KK.
DR WPI; 98-174915/16.
PT Production of equine chorionic gonadotropin hormone - having additional or modified sugar chains
PS Claim 2; Page 9; 14pp; Japanese.
CC This is the alpha-subunit of an equine chorionic gonadotropin (eCG) hormone. The eCG hormone is composed of alpha and beta-subunits and has an additional or modified sugar chain with follicle-stimulating hormone (FSH)-like and luteinising hormone (LH)-like activities. The eCG hormone can be stably produced in recombinant mammalian cells and can be used as an ovulation inducer and as an agent for the treatment of ovarian diseases.
CC Sequence 120 AA;
SQ
Query Match 53.1%; Score 78; DB 1; Length 120;
Best Local Similarity 53.6%; Pred. No. 0.00048;
Matches 15; Conservative 3; Mismatches 6; Indels 4; Gaps 1;
QY 2 pd-----vqdcpectlqenpfqpgapi 25
|| ||||| :|| :|| :||
Db 26 PDGEFTQDCPECKLRNKYFFKLGVP I 53
RESULT 26
W33815
ID W33815 standard; protein; 120 AA.
AC W33815;
DT 11-MAY-1998 (first entry)
DE Equine chorionic gonadotropin (eCG) alpha-subunit variant.
KW Equine; Chorionic gonadotropin; eCG; alpha-subunit; hormone; variant;
KW luteinising hormone; ovulation; treatment; ovarian disease.
OS Family Equidae.
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..24
FT /note= "signal peptide"
FT Protein 25..120
FT /note= "mature protein"
FT Misc_difference 56
FT /label= N56Q
FT /note= "wild type Asn is substituted by Gln"
PN J10036285-A.
PD 10-FEB-1998.
PF 23-JUL-1996; 193232.
PR 23-JUL-1996; JP-193232.
PA (ELED) DENKI KAGAKU KOGYO KK.
DR WPI; 98-174840/16.
PT Mutant equine chorionic gonadotropin hormone - has substitution in alpha chain preventing binding of sugar to residue 56, resulting in loss of luteinising hormone activity
PS Claim 3; Page -; 15pp; Japanese.
CC This is a variant of the alpha-subunit of equine chorionic gonadotropin (eCG). The wild type amino acid Asn at position 56 is replaced by Gln. This variant has an amino acid substitution to inhibit binding of a sugar chain to 56th amino acid resulting in loss of luteinising hormone (LH) activity. A variant gonadotropin hormone composed of the variant alpha-subunit and beta-subunit of eCG is a safe and widely applicable ovulation inducer and may be used as an agent for the treatment of ovarian diseases. This gonadotropin hormone is substantially free from LH activity.
CC Note: This sequence is not provided in the specification; it has been created from the wild type eCG alpha-subunit sequence provided in Page 9.
CC Sequence 120 AA;
SQ
Query Match 53.1%; Score 78; DB 1; Length 120;
Best Local Similarity 53.6%; Pred. No. 0.00048;
Matches 15; Conservative 3; Mismatches 6; Indels 4; Gaps 1;
QY 2 pd-----vqdcpectlqenpfqpgapi 25
|| ||||| :|| :|| :||
Db 26 PDGEFTQDCPECKLRNKYFFKLGVP I 53

QY 2 pd-----vqdcpectlqenpfqpgapi 25
|| ||||| :|| :|| :||
Db 26 PDGEFTQDCPECKLRNKYFFKLGVP I 53
RESULT 27
W33772
ID W33772 standard; protein; 120 AA.
AC W33772;
DT 11-MAY-1998 (first entry)
DE Equine chorionic gonadotropin (eCG) alpha-subunit.
KW Equine; Chorionic gonadotropin; eCG; alpha-subunit; hormone; variant;
KW luteinising hormone; ovulation; treatment; ovarian disease.
OS Family Equidae.
OS Location/Qualifiers
FH Key
FT Peptide 1..24
FT /note= "signal peptide"
FT Protein 25..120
FT /note= "mature protein"
FT Misc_difference 56
FT /note= "If amino acid at this position is Asn, then amino acid at position 58 is any other amino acid except Thr or Ser; if at this position is an amino acid any other than Asn, then at position 58, the amino acid is Thr or Ser; if amino acid at this position is Gln, then amino acid at position 58 is Thr"
FT Misc_difference 58
FT /note= "Amino acid at this position is Thr or Ser if amino acid at position 56 is not Asn; amino acid at this position is not Thr or Ser if amino acid at position 56 is Asn; amino acid at this position is Thr if amino acid at position 56 is Gln"
PN J10036285-A.
PD 10-FEB-1998.
PF 23-JUL-1996; 193232.
PR 23-JUL-1996; JP-193232.
PA (ELED) DENKI KAGAKU KOGYO KK.
DR WPI; 98-174840/16.
PT Mutant equine chorionic gonadotropin hormone - has substitution in alpha chain preventing binding of sugar to residue 56, resulting in loss of luteinising hormone activity
PS Claim 2; Page 9; 15pp; Japanese.
CC This is the alpha-subunit of equine chorionic gonadotropin (eCG). A variant of this alpha-subunit has an amino acid substitution to inhibit binding of a sugar chain to 56th amino acid resulting in loss of luteinising hormone (LH) activity. A gonadotropin hormone composed of the variant alpha-subunit and beta-subunit of eCG is a safe and widely applicable ovulation inducer and may be used as an agent for the treatment of ovarian diseases. This gonadotropin hormone is substantially free from LH activity.
CC Sequence 120 AA;
SQ
Query Match 53.1%; Score 78; DB 1; Length 120;
Best Local Similarity 53.6%; Pred. No. 0.00048;
Matches 15; Conservative 3; Mismatches 6; Indels 4; Gaps 1;
QY 2 pd-----vqdcpectlqenpfqpgapi 25
|| ||||| :|| :|| :||
Db 26 PDGEFTQDCPECKLRNKYFFKLGVP I 53
RESULT 28
W65109
ID W65109 standard; Protein; 120 AA.
AC W65109;
DT 28-SEP-1998 (first entry)
DE Equine chorionic gonadotropin alpha-chain protein.
KW Chorionic gonadotropin alpha chain; equine; eCG; luteinising hormone; LH; follicle-stimulating hormone; FSH; treatment; fertility; disorder;
KW

KW domestic animal; farm animal; multiple ovulation; ovulation inhibition;
 KW ovarian tumour.
 OS Equus caballus.
 FH Key
 FT Peptide
 FT 1. .24 Location/Qualifiers
 FT 25. .120 /label= signal
 FT Protein
 FT /label= chorionic gonadotropin alpha chain
 PN WO9821238-A1.
 PD 22-MAY-1998.
 PF 12-NOV-1997; J04113.
 PR 12-NOV-1996; JP-300041.
 PA (TEIK) TEIKOKU HORMONE MFG CO LTD.
 PI Min K, Ogawa T, Shiota K;
 DR WPI: 98-297865/26.
 DR N-PSDB; V35321.
 PT Recombinant single-stranded equine chorionic gonadotropin - has
 PT greatly increased FSH-type activity over the naturally occurring
 PT hormone
 PS Disclosure; Page 11; 26pp; Japanese.
 CC This sequence represents the equine chorionic gonadotropin (eCG) alpha
 CC chain. This sequence is used in a method which results in a recombinant
 CC eCG having the alpha-chain fused to the C-terminal side of the
 CC beta-chain. The encoded polypeptide has similar luteinising hormone (LH)
 CC activity to the naturally occurring or wild-type eCG but about seven
 CC times greater than follicle-stimulating hormone (FSH) activity. The
 CC polypeptide is used in the treatment of fertility disorders in animals
 CC including domestic animals (e.g. rat, mouse, rabbit) and farm animals
 CC (e.g. cow, pig, sheep or horse), such as multiple ovulation, ovulation
 CC inhibition, ovarian tumours, the timing of coming on heat after
 CC parturition, and decline of seminal gland functionality.
 SQ Sequence 120 AA;

Query Match 53.1%; Score 78; DB 1; Length 120;
 Best Local Similarity 53.6%; Pred. No. 0.00048;
 Matches 15; Conservative 3; Mismatches 6; Indels 4; Gaps 1;
 QY 2 pd----vqdcpecltqenpfssqpgapi 25
 DB 26 PDGEFTTQDCPECKLRNKYFFKLGVP I 53

RESULT 29
 ID W65111 standard; Protein; 265 AA.
 AC W65111;
 DT 28-SEP-1998 (first entry)
 DE Equine chorionic gonadotropin alpha and beta chain fusion protein.
 KW Chorionic gonadotropin beta chain; equine; eCG; luteinising hormone;
 KW LH; follicle-stimulating hormone; FSH; treatment; fertility; disorder;
 KW domestic animal; farm animal; multiple ovulation; ovulation inhibition;
 KW ovarian tumour.
 OS Equus caballus.
 FH Key
 FT Peptide
 FT 1. .19 Location/Qualifiers
 FT 20. .265 /label= signal
 FT Protein /label= alpha and beta fusion protein
 FT 20. .169 /label= beta-chain
 FT Region /label= beta-chain
 FT Misc_difference 107
 FT /note= "encoded by ATG"
 FT Region 170. .265
 FT /label= alpha-chain
 PN WO9821238-A1.
 PD 22-MAY-1998.
 PF 12-NOV-1997; J04113.
 PR 12-NOV-1996; JP-300041.
 PA (TEIK) TEIKOKU HORMONE MFG CO LTD.
 PI Min K, Ogawa T, Shiota K;
 DR WPI: 98-297865/26.
 DR N-PSDB; V35323.

PT Recombinant single-stranded equine chorionic gonadotropin - has
 PT greatly increased FSH-type activity over the naturally occurring
 PT hormone
 PS Claim 1; Page 14-15; 26pp; Japanese.
 CC This sequence represents a fusion protein composed of the equine
 CC chorionic gonadotropin (eCG) alpha chain fused to the C-terminal side of
 CC the beta-chain. The encoded polypeptide has similar luteinising hormone
 CC (LH) activity to the naturally occurring or wild-type eCG but about seven
 CC times greater than follicle-stimulating hormone (FSH) activity. The
 CC polypeptide is used in the treatment of fertility disorders in animals
 CC including domestic animals (e.g. rat, mouse, rabbit) and farm animals
 CC (e.g. cow, pig, sheep or horse), such as multiple ovulation, ovulation
 CC inhibition, ovarian tumours, the timing of coming on heat after
 CC parturition, and decline of seminal gland functionality.
 SQ Sequence 265 AA;

Query Match 53.1%; Score 78; DB 1; Length 265;
 Best Local Similarity 53.6%; Pred. No. 0.0011;
 Matches 15; Conservative 3; Mismatches 6; Indels 4; Gaps 1;
 QY 2 pd----vqdcpecltqenpfssqpgapi 25
 DB 171 PDGEFTTQDCPECKLRNKYFFKLGVP I 198

RESULT 30
 P50638
 ID P50638 standard; Protein; 89 AA.
 AC P50638;
 DT 15-JAN-1992 (first entry)
 DE Salmon gonadotrophin.
 KW Salmon hypophysis cerebri hormones; ACTH; gamma-LPH; beta-MSH; CLIP;
 KW Beta-endorphin; salmon gonadotrophin; E.coli.
 OS Oncorhynchus keta.
 PN J60176588-A.
 PD 10-SEP-1985.
 PR 24-FEB-1984; 032700.
 PA (SEKG) SEIKAGAKU KOGYO KK.
 DR WPI: 85-261180/42.
 DR N-PSDB; N50521.
 PT DNA plasmid and its prepn. - having deoxyribonucleotide
 PT sequence to code salmon-hypophysis cerebri-hormone (precursor).
 PS Disclosure; Page 562; 9pp; Japanese.
 CC The gene product may be expressed from an E.coli plasmid expression
 CC system for the mass production of various kinds of salmon hypophysis
 CC cerebri hormones eg. ACTH; beta-,gamma-LPH; alpha-,beta-MSH; CLIP;
 CC beta endorphin etc.
 SQ Sequence 89 AA;

Query Match 49.0%; Score 72; DB 1; Length 89;
 Best Local Similarity 63.2%; Pred. No. 0.0024;
 Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 7 cpecltqenpfssqpgapi 25
 DB 5 CEECKLRNKYFSPNGAPV 23

RESULT 31
 R54671
 ID R54671 standard; Protein; 119 AA.
 AC R54671;
 DT 10-JAN-1995 (first entry)
 DE Fish gonadotropic hormone alpha I chain.
 KW Polymerase chain reaction; PCR; amplify; primer; alpha; beta I;
 KW beta II; chain; fish; gonadotropin hormone; GTH; oviposition;
 KW animal; promotion of maturation.
 OS Oncorhynchus gorbuscha.
 PN J06107689-A.
 PD 19-APR-1994.

PF 20-NOV-1991; 354152.
 PR 20-NOV-1991; JP-354152.
 PA (NIOC) NIPPON OIL CO LTD.
 DR WPI; 94-163941/20.
 DR N-PSDB; Q65373.
 PT Fish gonadotropic hormone - and transformants producing it,
 PT useful for maturation and oviposition promotion in fish
 PS Claim 1; Page 8; 22pp; Japanese.
 CC The sequences given in R54671-74 represent the alpha and beta chains
 CC of the fish gonadotropin hormone (GTH). The DNA encoding these
 CC sequences are produced by PCR using DNA extracted from postovipository
 CC pink salmon trout pituitary. GTH may be produced by transforming a
 CC host organism with these sequences and collecting the protein from the
 CC culture medium. The GTH produced may be used for the promotion of
 CC maturation and oviposition in fish and other animals.
 SQ Sequence 119 AA;

Query Match 49.0%; Score 72; DB 1; Length 119;
 Best Local Similarity 63.2%; Pred. No. 0.0033;
 Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 7 cpeclqenpfsggapi 25
 DB 35 CEECKLNENKLFSPGAPV 53

RESULT 32
 W41794
 ID W41794 standard; peptide; 12 AA.
 AC W41794;
 DT 14-MAY-1998 (first entry)
 DE Human TSH alpha-subunit mutant Gln3Lys.
 KW Human; thyroid stimulating hormone; TSH; alpha-subunit;
 KW glycoprotein hormone; GPH; super-agonist; antagonist; mutant.
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc_difference 4
 FT /label= Q13K
 FT /note= "wild type Gln replaced by Lys"

W09742322-AL.
 13-NOV-1997.
 08-MAY-1996; U06483.
 08-MAY-1996; WO-U06483.
 (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Grossman M, Szkudlinski MW, Weintraub BD;
 WPI: 97-558984/51.
 PT New modified glyco-protein hormones - having amino acid
 PT substitutions from one species in another species to produce
 PT products having super-agonist or antagonist activity
 PS Example; Fig 1; 90pp; English.
 CC The present sequence, a human thyroid stimulating hormone (TSH)
 CC alpha-subunit mutant, was used in a novel method
 CC for the preparation of a human glycoprotein hormone (GPH),
 CC comprising a basic amino acid in the alpha-subunit at positions 11,
 CC 13, 14, 16, 17, and 20.
 CC The method can be used to obtain GPH having super-agonist or
 CC antagonist activity. The GPH agonists can be used to treat, e.g.
 CC Grave's disease, thyroid cancer, ovulatory dysfunction, luteal
 CC phase defect, unexplained infertility, male factor infertility and
 CC time-limited conception. They can also be used in the
 CC super-ovulation of bovine animals. The modified GPH can also be
 CC used to target delivery of therapeutic agents to thyroid or gonadal
 CC tissue, or in the treatment of certain neoplasms. The modified GPH
 CC can have an activity which is increased by at least 3-fold compared
 CC to wild-type GPH.
 SQ Sequence 12 AA;

Query Match 44.2%; Score 65; DB 1; Length 12;
 Best Local Similarity 91.7%; Pred. No. 0.0028;
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 ctlqenpfssqp 21
 DB 1 CTLKENPFSSQP 12

RESULT 33
 W41796
 ID W41796 standard; peptide; 12 AA.
 AC W41796;
 DT 14-MAY-1998 (first entry)
 DE Human TSH alpha-subunit mutant Gln20Lys.
 KW Human; thyroid stimulating hormone; TSH; alpha-subunit;
 KW glycoprotein hormone; GPH; super-agonist; antagonist; mutant.
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc_difference 11
 FT /label= Q20K
 FT /note= "wild type Gln replaced by Lys"

W09742322-AL.
 13-NOV-1997.
 08-MAY-1996; U06483.
 08-MAY-1996; WO-U06483.
 (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Grossman M, Szkudlinski MW, Weintraub BD;
 WPI: 97-558984/51.
 PT New modified glyco-protein hormones - having amino acid
 PT substitutions from one species in another species to produce
 PT products having super-agonist or antagonist activity
 PS Example; Fig 1; 90pp; English.
 CC The present sequence, a human thyroid stimulating hormone (TSH)
 CC alpha-subunit mutant, was used in a novel method
 CC for the preparation of a human glycoprotein hormone (GPH),
 CC comprising a basic amino acid in the alpha-subunit at positions 11,
 CC 13, 14, 16, 17, and 20.
 CC The method can be used to obtain GPH having super-agonist or
 CC antagonist activity. The GPH agonists can be used to treat, e.g.
 CC Grave's disease, thyroid cancer, ovulatory dysfunction, luteal
 CC phase defect, unexplained infertility, male factor infertility and
 CC time-limited conception. They can also be used in the
 CC super-ovulation of bovine animals. The modified GPH can also be
 CC used to target delivery of therapeutic agents to thyroid or gonadal
 CC tissue, or in the treatment of certain neoplasms. The modified GPH
 CC can have an activity which is increased by at least 3-fold compared
 CC to wild-type GPH.
 SQ Sequence 12 AA;

Query Match 44.2%; Score 65; DB 1; Length 12;
 Best Local Similarity 91.7%; Pred. No. 0.0028;
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 ctlqenpfssqp 21
 DB 1 CTLKENPFSSQP 12

RESULT 34
 W41795
 ID W41795 standard; peptide; 12 AA.
 AC W41795;
 DT 14-MAY-1998 (first entry)
 DE Human TSH alpha-subunit mutant Prol6Lys.
 KW Human; thyroid stimulating hormone; TSH; alpha-subunit;
 KW glycoprotein hormone; GPH; super-agonist; antagonist; mutant.
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc_difference 7
 FT /label= P16K
 FT /note= "wild type Pro replaced by Lys"

Query Match 44.2%; Score 65; DB 1; Length 12;
 Best Local Similarity 91.7%; Pred. No. 0.0028;
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;


```

||||:|||||:|
Db 1 CTLKENKFFSKP 12

RESULT 37
R38710
ID R38710 standard; peptide; 46 AA.
AC R38710;
DT 10-DEC-1993 (first entry)
DE CS-1 and CS-2 peptides
KW Lymphocyte; receptor; C-terminal cell binding domain; CTCBD;
KW fibronectin; Hep2; CS; connecting segment; adhesion; inhibition;
KW endothelial cell; skin disease; tripeptide; leucine-aspartate-valine;
KW alpha-4 beta-1; integrin; immunosuppressant; autoimmune disorder;
KW allergy; asthma.
OS Synthetic.
FH Key Location/Qualifiers
FT peptide 1..25
FT /label= CS-1
FT peptide 22..46
FT /label= CS-2
FT peptide 1..13
FT /label= A13
FT peptide 14..25
FT /label= B12
FT WO9312809-A.
PN PD
PD 08-JUL-1993.
PF 23-DEC-1992; U11191.
PR 24-DEC-1991; US-814873.
PA (HUTC-) HUTCHINSON CANCER RES CENT FRED.
PI Wayne EA;
DI WPI: 93-227054/28.
PT Lymphocytes adherence inhibition to endothelial cells for
PT treating skin disease - by exposing to tri:peptide
PT leucine-aspartate-valine to bind to alpha-4 beta-1, for
PT immunosuppressant treating auto immune disorders, allergy and
PT asthma
PS Disclosure: Fig 9: 98pp; English.
CC The alpha4beta1 integrin is a lymphocyte receptor for the C-terminal
CC cell binding domain (CTCBD) of fibronectin which comprises adhesion
CC sites in Hep2 and a high affinity site, CS-1, in the type III
CC connecting segment or V (for variable) region. Using a series of
CC peptides derived from CS-1, the tripeptide Leu-Asp-Val (LDV) is
CC identified as a minimal peptide capable of supporting stable
CC lymphocyte or melanoma cell adhesion.
SQ Sequence 46 AA;

Query Match 33.7%; Score 49.5; DB 1; Length 46;
Best Local Similarity 42.9%; Pred. No. 1.7;
Matches 9; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 2 pdvqdcptclqlenpfssqpg 22
|::| | | | | | | |
Db 17 PEILDVPS-TVQKTPFVTHPG 36

RESULT 38
P90823
ID P90823 standard; peptide; 446 AA.
AC P90823;
DT 28-JUN-1990 (first entry)
DE Amino terminal of tryptic/catheptic 66KD and 33KD heparin-binding frag-
DE ments and carboxyl terminal of 31 kD fragment of fibronectin A chain
KW Fibronectin A chain; heparin binding; fragment; neurite extension;
KW cell adhesion.
FH Key Location/Qualifiers
FT peptide 257..275
FT /note="peptide I"
FT peptide 297..315
FT /note="peptide II"
FT peptide 297..312
FT /note="peptide III"
FT

peptide 305..311
FT /note="Peptide IV"
FT peptide 71..86
FT /note="Peptide V"
FT peptide 134..142
FT /note="Peptide VI"
FT peptide 243..250
FT /note="Peptide VII"
FT region 26..43
FT /note="Amino terminal sequence of tryptic/catheptic (t/c)
FT t/c33 and t/c66 heparin binding fragments of fibronectin"
FT region 391..400
FT /note="Portion of amino terminal of tryptic (t) t31
FT fragment of fibronectin"
FT region 432..443
FT /note="Portion of trypsin (t) 31 fragment of fibronectin"
FT region 390..390
FT /note="Carboxyl terminal limit of the t/c33 heparin
FT binding fragment"
FT WO8901942-A.
PN PD
PD 09-MAR-1989.
PF 24-AUG-1988; U02913.
PR 27-JUL-1988; US-225045, US-089073.
PA (MINU) Minnesota University.
PI Furcht LT, McCarthy JB;
DI WPI: 89-085521/11.
PT New polypeptide fibronectin fragments -
PT useful for promoting cell adhesion, heparin binding and/or
PT neurite extension
PS Figure 11; 45pp; English.
CC Residue 1 corresponds to residue 1558 of intact plasma fibronectin.
CC Peptides I-VII are fragments of the 33KD carboxy-terminal (t/c33)
CC heparin-binding region of fibronectin A chain. They promote neurite
CC extension, promote adhesion and spreading of endothelial and melanoma
CC cells and/or promote adhesion of heparin to synthetic substrates.
CC They may be useful for assisting nerve regeneration, promoting wound
CC healing and implant acceptance, promoting cell attachment to culture
CC substrates, inhibiting metastasis of malignant cells and/or binding
CC excess heparin in vivo in heparin therapy.
SQ Sequence 446 AA;

Query Match 33.7%; Score 49.5; DB 1; Length 446;
Best Local Similarity 42.9%; Pred. No. 18;
Matches 9; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 2 pdvqdcptclqlenpfssqpg 22
|::| | | | | | | |
Db 328 PEILDVPS-TVQKTPFVTHPG 347

RESULT 39
R92778
ID R92778 standard; Protein; 2324 AA.
AC R92778;
DT 21-JUN-1996 (first entry)
DE Human fibronectin.
KW Fibrin-binding peptide; fibronectin; fibrinolysis; fibrinogenesis;
KW thrombus; pulmonary embolus; atherosclerosis; tumour; diagnosis;
KW therapy.
OS Homo sapiens.
FH Key Location/Qualifiers
FT domain 21..65
FT /label= 1F1
FT /note= "type 1 module 1"
FT domain 66..109
FT /label= 2F1
FT /note= "type 1 module 2"
FT domain 110..154
FT /label= 3F1
FT /note= "type 1 module 3"
FT domain 155..199
FT /label= 4F1
FT

```



```
FT domain /note= "type 1 module 4"  
FT 200..244  
FT /label= 5F1  
FT /note= "type 1 module 5"  
FT 277..312  
FT /label= 6F1  
FT /note= "type 1 module 6"  
FT 439..486  
FT /label= 7F1  
FT /note= "type 1 module 7"  
FT 439..467  
FT /note= "Cys439-Cys467 disulfide bond"  
FT 465..477  
FT /note= "Cys465-Cys477 disulfide bond"  
FT 487..529  
FT /label= 8F1  
FT /note= "type 1 module 8"  
FT 530..569  
FT /label= 9F1  
FT /note= "type 1 module 9"  
FT 2123..2230  
FT /label= Fibrin_binding_domain  
FT /note= "11 kDa C-terminal fibrin-binding domain"  
FT 2141..2185  
FT /label= 10F1  
FT /note= "type 1 module 10"  
FT 2187..2230  
FT /label= 11F1  
FT /note= "type 1 module 11"  
FT 2144..2173  
FT /note= "Cys2144-Cys2173 disulfide bond"  
FT 2171..2183  
FT /note= "Cys2171-Cys2183 disulfide bond"  
FT 2233..2271  
FT /label= 12F1  
FT /note= "type 1 module 12"  
FT 2189..2216  
FT /note= "Cys2189-Cys2216 disulfide bond"  
FT 2214..2226  
FT /note= "Cys2214-Cys2226 disulfide bond"  
PN W09604304-A1.  
PD 15-FEB-1996.  
PF 01-AUG-1995; U09819.  
PR 01-AUG-1994; US-283857.  
PA (UNY ) UNIV NEW YORK STATE.  
PI Gold LI, Rostagno AA;  
DR WPI; 96-129333/13.  
PT New fibrin-binding peptide molecules - used for the diagnosis and  
PT treatment of conditions associated with fibrin deposition, e.g.  
PT thrombi  
PS Claim 2; Page 104-110; 146pp; English.  
CC Human fibronectin (R92778) has a fibrin-binding site, close to the C-  
CC terminus (amino acids 2123-2232 or 2141-2230), covering the 10F1.11F1  
CC module pair of the protein. Fibrin-binding sites can be prep. by  
CC enzymatic cleavage of fibronectin, peptide synthesis or by recombinant  
CC DNA techniques. They are used to detect a fibrin-binding target site,  
CC to treat disorders involving abnormal fibrinolysis or fibrinogenesis,  
CC to disrupt blood clots and to aid delivery of medicaments to fibrin-  
CC contg. sites. They show high affinity to and slow dissociation from  
CC fibrin, and provide fast diffusion and rapid clearance.  
SQ Sequence 2324 AA;
```

Query Match 33.7%; Score 49.5; DB 1; Length 2324;
Best Local Similarity 42.9%; Pred. No. 1e+02; Mismatches 6; Indels 1; Gaps 1;
Matches 9; Conservative 5;

```
Qy 2 pdvqdqcectqlqenpfssqpg 22  
|::| |::| |::| |::| |::|  
Db 1977 PEILDVPS-TVQKTPPVTHPG 1996
```

```
RESULT 40  
P70373  
ID P70373 standard; protein; 2327 AA.  
AC P70373;  
DT 11-MAR-1991 (first entry)  
DE Human fibronectin gene product.  
FN; collagen; fibrin; heparin.  
OS Homo sapiens.  
PN EP-207751-A.  
PD 07-JAN-1987; 304998.  
PF 27-JUN-1986; 304998.  
PR 28-JUN-1985; GB-016421.  
PA (DELTA-) DELTA BIOTECHN.  
PI Baralle FE;  
DR WPI; 87-001441/01.  
DR N-PSDB; N70596.  
PT New fibronectin polypeptide sequence with affinity for collagen  
PT etc. - useful for targeting therapeutic substances on natural  
PT fibrin, for use in affinity purificn. of polypeptide(s) etc.  
PS Claim 11; Fig 3A; 32pp; English.  
CC The product may be expressed from a transformed micro-organism,  
CC esp. E.coli.  
CC FN binds to fibrin, heparin and Staphylococcus aureus, and may be  
CC used to target a therapeutic agent onto natural fibrin eg. a blood  
CC clot. It may also be used in affinity purification of a polypeptide,  
CC conjugated to the collagen binding site of FN and immobilised on a  
CC collagen surface.  
SQ Sequence 2327 AA;
```

Query Match 33.7%; Score 49.5; DB 1; Length 2327;
Best Local Similarity 42.9%; Pred. No. 1e+02; Mismatches 6; Indels 1; Gaps 1;
Matches 9; Conservative 5;

```
Qy 2 pdvqdqcectqlqenpfssqpg 22  
|::| |::| |::| |::| |::|  
Db 1980 PEILDVPS-TVQKTPPVTHPG 1999
```

```
RESULT 41  
R15468  
ID R15468 standard; Protein; 2327 AA.  
AC R15468;  
DT 12-MAR-1992 (first entry)  
DE Human fibronectin.  
KW Fibrin-imaging; atherosclerosis; thrombus inhibitor.  
OS Homo sapiens.  
PN W09117765-A.  
PD 28-NOV-1991.  
PF 21-MAY-1991; U03584.  
PR 21-MAY-1990; US-526397.  
PA (BIOT-) BIO-TECHN GEN CORP.  
PI Vogel T, Levanon A, Werber M, Guy R, Panet A, Hartman J,  
PI Shaked H;  
DR WPI; 91-369004/50.  
DR N-PSDB; Q15214.  
PT New fibrin binding domain polypeptide(s) - useful in imaging  
PT fibrin-contg. substances, to inhibit thrombus formation and treat  
PT wounds  
PS Disclosure: Fig 2; 191pp; English.  
CC The amino acid sequence is that of human fibronectin, this can be used  
CC to derive polypeptides which are identical to part of the fibrin-  
CC binding domain (FBD) of fibronectin. These polypeptides can be used  
CC to inhibit thrombus formation; or (coupled to a thrombolytic agent)  
CC to induce thrombolysis, or to treat wounds, e.g. in skin, eyes or  
CC tendons (in conjunction with a polypeptide which includes a part of  
CC the cell-binding domain (CBD) of fibronectin). These polypeptides are  
CC easier to prepare than the full 31kD polypeptide. It can also be  
CC used to image fibrin-contg. materials, esp. a thrombus or athero-  
CC sclerotic plaque, pref. using a gamma counter.  
SQ Sequence 2327 AA;
```



```
Query Match      33.7%; Score 49.5; DB 1; Length 2327;
Best Local Similarity 42.9%; Pred. No. 1.le+02;
Matches 9; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 2 pdvqdcpectqlqenpfssqpg 22
   |::| | |::| | |
Db 1980 PEILDVPS-TVQKTPFVTHPG 1999

RESULT 42
W63171
ID W63171 standard; peptide; 2386 AA.
AC W63171;
DE 27-OCT-1998 (first entry)
DT Amino acid sequence of fibronectin.
KW Mast cell protease; MCP; mouse; inhibitor; peptide substrate; asthma;
KW tryptase-6 protein; inflammatory disorder; allergic rhinitis; urticaria;
KW antioedema; eczematous dermatitis; atopic dermatitis; anaphylaxis;
KW hyperproliferative skin disease; peptic ulcer; hyperresponsiveness;
KW inflammatory skin condition; fibronectin.
OS Mus sp.
PN WO9833812-A1.
PD 06-AUG-1998.
PF 30-JAN-1998; U01865.
PR 05-FEB-1997; US-037090.
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL.
PI Huang C, Stevens RL;
DR WPI; 98-437390/37.
PT Tryptase-6 complex inhibitory peptides - used to treat mast
PT cell-mediated inflammatory disorders e.g. asthma
PS Disclosure; pages 36-42; 69pp; English.
CC This is the amino acid sequence of fibronectin. The tryptase-6 complex
CC inhibitory peptides of the invention comprise the sequence of a fragment
CC of the fibronectin (residues 1351-1356). Sequences shown in W63160 to
CC W63169 represent mouse mast cell protease (mMCP-6) inhibitory peptides.
CC These peptides which are tryptase-6 complex inhibitors, can be used for
CC treating a mast cell-mediated inflammatory disorder. The inhibitors can
CC be used to treat inflammatory disorders including asthma, allergic
CC rhinitis, urticaria and antioedema, eczematous dermatitis (atopic
CC dermatitis), anaphylaxis, hyperproliferative skin disease, peptic ulcers,
CC inflammatory bowel disorder, hyperresponsiveness and inflammatory skin
CC conditions.
SQ Sequence 2386 AA;

Query Match      33.7%; Score 49.5; DB 1; Length 2386;
Best Local Similarity 42.9%; Pred. No. 1.le+02;
Matches 9; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 2 pdvqdcpectqlqenpfssqpg 22
   |::| | |::| | |
Db 2008 PEILDVPS-TVQKTPFVTHPG 2027

RESULT 43
R60021
ID R60021 standard; Protein; 2446 AA.
AC R60021;
DE 23-FEB-1995 (first entry)
DT Fibrinogen-alpha.
KW Tissue binding; tissue sealing; wound healing; vulneryary;
KW tissue-binding domain; TBD; crosslinking domain; fibronectin;
KW fibrinogen-alpha; heparin-binding domain; collagen-binding domain;
KW cell-binding domain; hybrid protein.
OS Homo sapiens.
FH Key Location/Qualifiers
FT domain 282..608
FT /label= collagen-binding_domain
FT /note= "acts as tissue-binding domain of hybrid
FT protein"
FT 1812..2171
FT /label= heparin-binding_domain
FT /note= "acts as tissue-binding domain of hybrid
```

```
FT WO9416085-A.
PN 21-JUL-1994.
PD 30-DEC-1993; U12687.
PF 30-DEC-1992; US-998271.
PA (ZYMO ) ZYMOGENETICS INC.
PI Irani MH;
DR WPI; 94-249231/30.
DR N-PSDB; Q70009.
DE New hybrid proteins for use in tissue sealing and wound healing -
DE comprising a tissue-binding domain from a protein covalently
DE linked to a crosslinking domain of another protein
PS Disclosure; Page 37-48; 87pp; English.
CC Hybrid proteins have a tissue-binding domain (TBD) from 1 protein
CC linked to a crosslinking domain from another protein. The TBD
CC comprises: aa 2-946, 928-1338 and especially 2-1336 of the sequence
CC given in R60019; the heparin-binding domain (aa 1812-2171 of R60021)
CC of fibronectin; the collagen-binding domain (aa 282-608 of R60021)
CC of fibronectin; or the cell-binding domain (aa 1357-1903 or 1532-
CC 1631 of R60020) of fibronectin. DNA encoding a fibronectin-
CC fibrinogen hybrid is given in Q70007, and sequences for fibronectin
CC and fibrinogen-alpha in Q70008 and Q70009, respectively.
SQ Sequence 2446 AA;

Query Match      33.7%; Score 49.5; DB 1; Length 2446;
Best Local Similarity 42.9%; Pred. No. 1.le+02;
Matches 9; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 2 pdvqdcpectqlqenpfssqpg 22
   |::| | |::| | |
Db 2099 PEILDVPS-TVQKTPFVTHPG 2118

RESULT 44
W99595
ID W99595 standard; peptide; 2477 AA.
AC W99595;
DE 22-JUN-1999 (first entry)
DT Human fibronectin.
KW Human; fibronectin; glycoprotein; extracellular matrix; motif; migration;
KW stimulation; wound healing; periodontal tissue regeneration; metastasis;
KW tumour; invasion; angiogenesis; inflammation; connective tissue function.
OS Homo sapiens.
PN WO9902674-A1.
PD 21-JAN-1999.
PF 01-JUL-1998; G01939.
PR 08-JUL-1997; GB-014276.
PA (UYDU-) UNIV DUNDEE.
PI Schor SL;
DR WPI; 99-120875/10.
DE New peptides which modulate cell migration have the motif IGD - and
DE are useful to stimulate angiogenesis
PS Disclosure; Fig 5; 57pp; English.
CC This sequence represents the amino acid sequence of human fibronectin,
CC a widely distributed glycoprotein present in extracellular matrices.
CC The protein contains the motifs IGDS (W99593) and IGDQ (W99594) which
CC can be used in peptides to stimulate cell migration for wound healing,
CC periodontal tissue regeneration, angiogenesis, inhibition of tumour
CC invasion and metastasis, and in relation to inflammation or connective
CC tissue function.
SQ Sequence 2477 AA;

Query Match      33.7%; Score 49.5; DB 1; Length 2477;
Best Local Similarity 42.9%; Pred. No. 1.le+02;
Matches 9; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 2 pdvqdcpectqlqenpfssqpg 22
   |::| | |::| | |
Db 2099 PEILDVPS-TVQKTPFVTHPG 2118
```



```
RESULT 45
W41799
ID W41799 standard; peptide; 12 AA.
AC W41799;
DE Human TSH alpha-subunit mutant Pro16Lys/Gln20Lys/Gln13Lys/Glu14Lys.
KW Human; thyroid stimulating hormone; TSH; alpha-subunit;
KW glycoprotein hormone; GPH; super-agonist; antagonist; mutant.
OS Homo sapiens.
OS Synthetic.
FH Key
FT Misc_difference 7 Location/Qualifiers
FT /label= P16K
FT /note= "wild type Pro replaced by Lys"
FT Misc_difference 11
FT /label= Q20K
FT /note= "wild type Gln replaced by Lys"
FT Misc_difference 4
FT /label= Q13K
FT /note= "wild type Gln replaced by Lys"
FT Misc_difference 5
FT /label= E14K
FT /note= "wild type Glu replaced by Lys"
FT
FT WO9742322-AL.
PN 13-NOV-1997.
PD 08-MAY-1996; U06483.
PF 08-MAY-1996; WO-U06483.
PR (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI Grossman M, Szkudlinski MW, Weintraub BD;
PI WPI; 97-558984/51.
PT New modified glyco-protein hormones - having amino acid
PT substitutions from one species in another species to produce
PT products having super-agonist or antagonist activity
PT Example; Fig 1; 90pp; English.
PS The present sequence, a human thyroid stimulating hormone (TSH)
CC alpha-subunit mutant, was used in a novel method
CC for the preparation of a human glycoprotein hormone (GPH),
CC comprising a basic amino acid in the alpha-subunit at positions 11,
CC 13, 14, 16, 17, and 20.
CC The method can be used to obtain GPH having super-agonist or
CC antagonist activity. The GPH agonists can be used to treat, e.g.
CC Grave's disease, thyroid cancer, ovulatory dysfunction, luteal
CC phase defect, unexplained infertility, male factor infertility and
CC time-limited conception. They can also be used in the
CC super-ovulation of bovine animals. The modified GPH can also be
CC used to target delivery of therapeutic agents to thyroid or gonadal
CC tissue, or in the treatment of certain neoplasms. The modified GPH
CC can have an activity which is increased by at least 3-fold compared
CC to wild-type GPH.
CC Sequence 12 AA;
SQ

Query Match 33.3%; Score 49; DB 1; Length 12;
Best Local Similarity 66.7%; Pred. No. 0.49;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 10 cttlqenpfssq 21
Db 1 CTLKKNFFSKP 12

RESULT 46
W37832
ID W37832 standard; Protein; 297 AA.
AC W37832;
DE Human adrenocorticotrophic stimulating hormone receptor.
KW Human adrenocorticotrophic stimulating hormone receptor gene;
KW metabolic disorder; melanocortin receptor; MCR; agonist; antagonist;
KW feeding; eating disorder; anorexia; obesity; cachexia; cancer;
KW inhibition.
OS Homo sapiens.
PN WO9810068-A2.

Query Match 33.3%; Score 49; DB 1; Length 12;
Best Local Similarity 66.7%; Pred. No. 0.49;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 10 cttlqenpfssq 21
Db 1 CTLKKNFFSKP 12
```

```
PD 12-MAR-1998.
PF 04-SEP-1997; U15565.
PR 04-SEP-1996; US-706281.
PA (UYOR-) UNIV OREGON HEALTH SCI.
PI Boston BA, Chen W, Cone RD, Fan W, Kesterton RA,
PI Lu D;
DR WPI; 98-193618/17.
DR N-PSDB: V19137.
PT Identifying melanocortin receptor agonists and antagonists - using a
PT panel of recombinant mammalian cells expressing alpha-melanocyte
PT stimulating hormone, ACTH, MC-3, MC-4 and MC-5 receptors
PS Example 2C; Fig 3A-B; 121pp; English.
CC This is the amino acid sequence of the human adrenocorticotrophic
CC stimulating hormone receptor (melanocyte stimulating hormone-2 receptor
CC (MCR2)). The MCR agonists and antagonists can be used for modifying
CC feeding behaviour in an animal. The antagonists can be used for
CC stimulating feeding while the agonists can be used for inhibiting
CC feeding. They can be used for the treatment of eating disorders such
CC as anorexia and obesity, and other pathological weight and
CC eating-related disorders. They can also be used to treat failure
CC to thrive disorders and disease-related cachexia, such as occurs in
CC cancer patients, as well as other metabolic disorders.
SQ Sequence 297 AA;

Query Match 33.3%; Score 49; DB 1; Length 297;
Best Local Similarity 57.1%; Pred. No. 14;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 6 dcpectlqenpffs 19
Db 20 DCPRCVLPPEIFFT 33

RESULT 47
W87866
ID W87866 standard; Protein; 297 AA.
AC W87866;
DE Human melanocortin receptor MC2-R.
KW Melanocortin receptor; MC2-R; alpha-melanocyte stimulating hormone;
KW MSH; human; acne; therapy.
OS Mus sp.
PN WO9855914-AL.
PD 17-DEC-1998.
PF 12-JUN-1998; U12098.
PF 13-JUN-1997; US-050063.
PA (UYOR-) UNIV OREGON HEALTH SCI.
PI Chen W, Cone RD, Low MJ;
PI WPI; 99-080902/07.
DR N-PSDB: V63704.
PT Identifying compounds that bind to melanocortin receptors - such as
PT therapeutic agents for treating exocrine disorders like acne
PS Example 2C; Page 60-61; 144pp; English.
CC This is the amino acid sequence of human melanocortin receptor
CC MC2-R. The sequence was deduced from the nucleotide sequence of a
CC genomic DNA clone (see V63704). MC2-R is an adrenocorticotrophic
CC hormone (ACTH) receptor. The invention relates to the cloning,
CC expression and functional characterisation of mammalian
CC melanocortin receptor MC1-R, MC2-R, MC3-R, MC4-R and MC5-R nucleic
CC acids (see V63702-08) and polypeptides (see W87864-70), as well as
CC expression constructs, eukaryotic cells transformed with such
CC constructs, knockout animals, and methods and reagents for
CC developing agonists and antagonists specific for mammalian
CC melanocortin receptors. Such compounds, particularly those
CC specific for MC5-R, are used to treat disorders of exocrine gland
CC function, e.g. of the lacrimal or sebaceous glands, particularly
CC acne, other skin disorders and 'dry eye', also disorders related to
CC oestrus, mating, gestation and other pheromone-related conditions.
SQ Sequence 297 AA;

Query Match 33.3%; Score 49; DB 1; Length 297;
```


Best Local Similarity 57.1%; Pred. No. 14;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 6 dpectqlqenpffs 19
| | | | | | | |
DB 20 DPCRVLPPEIFFT 33

RESULT 48

ID W60611 standard; Protein; 313 AA.
AC W60611;
DT 28-SEP-1998 (first entry)
DE Human mucosal addressin cell adhesion molecule-1(e).
KW Mucosal addressin cell adhesion molecule-1; MadCAM-1(e); human;
cancer; tumour; inflammation; transplant rejection; arthritis;
rheumatoid arthritis; infection; inflammatory bowel disease;
autoimmune disease; experimental autoimmune encephalitis;
dermatosis; diagnosis; therapy.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..17
FT /label= Sig_peptide
FT Protein 18..313
FT /label= Mat_protein
FT /note= "Claim 16"
FT Domain 18..250
FT /label= Extracellular
FT /note= "Claim 16"
FT Domain 251..270
FT /label= Transmembrane
FT /note= "Claim 16"
FT Domain 271..313
FT /label= Intracellular
FT /note= "Claim 16"
PN W09820110-A1.
PD 14-MAY-1998.
PF 01-NOV-1996; U17549.
PR (HUMA-) HUMAN GENOME SCI INC.
PA (UYAU-) UNIV AUCLAND.
PI Greene JM, Krissansen GW, Leung EYF, Ni J, Ruben SM;
DR WPI; 98-286926/25.
DR N-PSDB; V38205.
PT Addressin cell adhesion molecules - used to develop products for
detection of inflammatory conditions or cancer and for treating or
preventing inflammatory conditions
PS Claim 16; Page 102-104; 164pp; English.
CC This polypeptide comprises human mucosal vascular addressin cell
adhesion molecule MadCAM-1(e), a novel cell surface adhesion
molecule that shows homology to murine MadCAM-1. The invention
relates to human MadCAM-1(a) as well as 4 splice variants,
CC designated MadCAM-1(b), -1(c), -1(d) and -1(e) (see W60607-11).
CC Their amino acid sequences were deduced from cDNA clones (see
V38201-05) isolated from a human foetal brain cell cDNA library.
CC The invention also provides vectors, host cells, recombinant
methods of producing the polypeptides, as well as methods for
identifying agonists and antagonists of activity, diagnostic
methods for detecting cancer or a pathological inflammatory
condition, and therapeutic methods for treating an individual in
need of a reduction in activity of any of MadCAM-1(a-e). Genomic
DNA molecules comprising the 5' untranslated region and exons 1-5
CC (see V38223-28) which, in various combinations, comprise the coding
CC region of any of the MadCAM-1 splice variants are also claimed.
CC The novel human MadCAM-1 polypeptides can be used as a target for
the diagnosis and treatment of inflammatory conditions such as
transplant rejection, arthritis, rheumatoid arthritis, infection,
dermatosis, inflammatory bowel disease, and autoimmune disease,
CC including chronic relapsing experimental autoimmune encephalitis.
SQ Sequence 313 AA;

Query Match 33.3%; Score 49; DB 1; Length 313;

Best Local Similarity 39.1%; Pred. No. 15;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 pdvdcpectqlqenpffsqpgap 24
| | | | | | | | | |
DB 166 PEVQESPDTSPEPDTTSPEPP 188

RESULT 49

ID R86258 standard; Protein; 181 AA.
AC R86258;
DT 02-MAY-1996 (first entry)
DE Human CG beta-subunit (N130) lacking first glycosylation site.
KW Single chain gonadotropin; human chorionic gonadotropin; hCG;
alpha; beta; subunit; analogue; glycoprotein hormone; fertility;
inhibit; stimulate; increase; lutropin; luteinising hormone; LH;
KW follicle stimulating hormone; FSH; vaccine; contraceptive;
KW deglycosylated; glycosylation site; deletion.
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..20
FT /label= leader
FT region 21..165
FT /label= hCG_beta_subunit_(1-145)
FT /note= "the native glycosylation site at position
13 of hCG beta has been removed"
FT misc_difference 33
FT /note= "wild-type Asn 13 has been replaced by Gln
to remove a glycosylation site"
FT misc_difference 70
FT /note= "Arg corresponds to CCG codon"
FT region 166..173
FT /label= linker
FT region 174..181
FT /label= Gonadotropin_alpha_subunit_(1-8)
FT /note= "first 8 residues of alpha subunit"
PN W09522340-A1.
PD 24-AUG-1995.
PF 17-FEB-1995; U02067.
PR 18-FEB-1994; US-199382.
PA (SENS-) SENS-TEST.
PI Moyle WR;
DR WPI; 95-302553/39.
DR N-PSDB; T03240.
PT Methods for altering fertility in mammals, esp. humans - e.g.
stimulating fertility by reducing the activity and/or levels of
circulating glyco:protein hormones having lutropin activity
PS Example 23; Fig 17; 102pp; English.
CC Single chain gonadotropins having a chorionic gonadotropin (CG)
beta-subunit at the N-terminus and a CG alpha-subunit at the
C-terminus (or vice versa), joined by a linker of 1-16 amino acids
CC are claimed. Ten specific analogues were synthesised, each of which
CC was expected to contain 4 Asn-linked oligosaccharides. Removal of
CC the glycosylation sites, particularly from the alpha-subunit, has
CC been shown to reduce hormone efficacy and inhibit signal
transduction. The present sequence is that of the hCG beta-subunit
CC from which one of the two native glycosylation sites has been
removed using strand overlap PCR.
SQ Sequence 181 AA;

Query Match 32.7%; Score 48; DB 1; Length 181;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 apdvqgdep 8
| | | | | | | |
DB 174 APDVQDCP 181

RESULT 50
W13826

ID W13826 standard; Protein; 1420 AA.
 AC W13826;
 DT 04-JUN-1997 (first entry)
 DE yeast transcription regulatory factor SRB9.
 KW transcription regulatory factor; suppressor of RNA polymerase B;
 KW SRB9; RNA polymerase II; holoenzyme; SWI/SNF.
 OS Saccharomyces cerevisiae.
 PN W09708301-A1.
 PD 06-MAR-1997.
 PF 28-AUG-1996; U14192.
 PR 31-AUG-1995; US-521872.
 PR 11-OCT-1995; US-540804.
 PR 26-JAN-1996; US-590399.
 PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
 PI Chao DM, Koleske AJ, Thompson CW, Young RA;
 DR WPI; 97-179258/16.
 DR N-PSDB; T59909.
 PT Purified RNA polymerase II holoenzyme - comprises RNA polymerase II
 PT and one or more regulatory proteins, pref. suppressor of RNA
 PT polymerase B proteins or SWI/SNF proteins
 PS Claim II; Fig 11a-c; 154pp; English.
 CC Novel yeast SRB (suppressor of RNA polymerase B) proteins SRB4, SRB5,
 CC SRB6, SRB7, SRB8, SRB9, SRB10 and SRB11 (W13821-28) are transcription
 CC regulatory factors that act as positive and negative regulators of
 CC RNA polymerase II activity, and are components of the RNA polymerase
 CC II holoenzyme. They were identified using methods designed to
 CC identify transcription factors involved in RNA polymerase II
 CC C-terminal domain (CTD) function. SRB8 and SRB9 appear to repress
 CC CTD activity. Genomic clones (T59904-11) for the SRBs have been obtd.
 CC SRBs can be used to treat diseases resulting from alteration or
 CC deletion of the SRB gene, pref. by gene transfer technology. They
 CC can also be used in in vitro transcription of DNA and to identify
 CC cpds. that modify gene transcription.
 SQ Sequence 1420 AA;

Query Match 32.7%; Score 48; DB 1; Length 1420;
 Best Local Similarity 45.5%; Pred. NO. ie+02;
 Matches 10; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

Qy 6 dcp--ectlqenpfqpgapi 25

Db 576 DIPIDMTLPTSLYWDPGAPL 597

Search completed: July 14, 2000, 09:31:11
 Job time: 2128 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 14, 2000, 09:31:52 ; Search time 35.12 Seconds
(without alignments)
10.688 Million cell updates/sec

Title: ALPHA-CHAIN

Perfect score: 147

Sequence: 1 apdvcdpctqlqenpfsgapil 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 145308 seqs, 14437401 residues

Total number of hits satisfying chosen parameters: 145306

Minimum DB seq length: 0

Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%

Maximum Match 99%

Listing first 1000 summaries

Database :

Issued Patents AA:*

1: /cgnl_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgnl_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgnl_6/ptodata/1/iaa/6_COMB.pep.*

4: /cgnl_6/ptodata/1/iaa/PCTUS_COMB.pep.*

5: /cgnl_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49.5	33.7	46	US-08-338-282-1	Sequence 1, Appli
2	49.5	33.7	2231	US-08-153-799-16	Sequence 16, Appli
3	49.5	33.7	2324	US-08-283-857-1	Sequence 1, Appli
4	49.5	33.7	2324	PCT-US95-09819-1	Sequence 1, Appli
5	49.5	33.7	2327	5455158-1	Patent No. 5455158
6	49.5	33.7	2386	US-09-016-366A-12	Sequence 12, Appli
7	49.5	33.7	2446	US-08-551-356-2	Sequence 2, Appli
8	49.5	33.7	2446	PCT-US93-12687-2	Sequence 2, Appli
9	48	32.7	1420	US-08-540-804-14	Sequence 14, Appli
10	48	32.7	1420	US-08-218-265-14	Sequence 14, Appli
11	48	32.7	1420	US-08-521-872-14	Sequence 14, Appli
12	45	30.6	112	US-08-425-673-6	Sequence 6, Appli
13	44.5	30.3	530	US-08-513-841-1	Sequence 1, Appli
14	44.5	30.3	530	US-08-696-834-1	Sequence 1, Appli
15	44.5	30.3	530	US-08-942-673-1	Sequence 1, Appli
16	44.5	30.3	1013	US-08-866-650-5	Sequence 5, Appli
17	44.5	30.3	1013	US-09-021-287-5	Sequence 5, Appli
18	44.5	30.3	1013	US-08-991-408-2	Sequence 2, Appli
19	44	29.9	871	US-08-775-009-34	Sequence 34, Appli
20	44	29.9	1711	US-08-342-930-2	Sequence 2, Appli
21	44	29.9	4544	US-08-469-486-52	Sequence 52, Appli
22	44	29.9	4544	US-08-469-658-52	Sequence 52, Appli
23	43.5	29.6	402	US-08-709-979A-3	Sequence 3, Appli
24	43.5	29.6	415	US-08-833-642A-5	Sequence 5, Appli
25	43.5	29.6	435	US-08-361-920-27	Sequence 27, Appli
26	43.5	29.6	435	US-08-479-939-27	Sequence 27, Appli
27	43.5	29.6	435	US-08-483-432-27	Sequence 27, Appli
28	43	29.3	1694	US-08-494-168-2	Sequence 2, Appli

29	42	28.6	80	3	US-08-813-884-42	Sequence 42, Appli
30	42	28.6	369	2	US-08-991-300-2	Sequence 2, Appli
31	42	28.6	1093	3	US-08-545-860D-55	Sequence 55, Appli
32	42	28.6	1093	4	PCT-US94-04496-55	Sequence 55, Appli
33	41	27.9	454	1	US-08-166-316-2	Sequence 2, Appli
34	41	27.9	552	4	PCT-US93-07832-23	Sequence 23, Appli
35	41	27.9	787	2	US-08-720-484A-4	Sequence 4, Appli
36	40.5	27.6	1013	2	US-08-866-650-3	Sequence 3, Appli
37	40.5	27.6	1013	2	US-09-021-287-3	Sequence 3, Appli
38	40.5	27.6	1210	2	US-08-484-438-7	Sequence 7, Appli
39	40.5	27.6	1210	2	US-08-475-035-4	Sequence 4, Appli
40	40.5	27.6	1323	1	US-08-026-138E-4	Sequence 4, Appli
41	40.5	27.6	1336	2	US-08-231-193A-58	Sequence 58, Appli
42	40.5	27.6	1336	2	US-08-486-273A-58	Sequence 58, Appli
43	40	27.2	263	2	US-07-857-224B-13	Sequence 13, Appli
44	40	27.2	264	2	US-07-857-224B-11	Sequence 11, Appli
45	40	27.2	264	2	US-07-857-224B-12	Sequence 12, Appli
46	40	27.2	289	2	US-08-741-437-1	Sequence 1, Appli
47	40	27.2	289	2	US-09-134-593-1	Sequence 1, Appli
48	40	27.2	671	5	5266464-2	Patent No. 5266464
49	40	27.2	1021	1	US-08-497-025-3	Sequence 3, Appli
50	40	27.2	1848	4	PCT-US95-10661A-6	Sequence 6, Appli
51	40	27.2	4303	2	US-08-460-751-2	Sequence 2, Appli
52	39.5	26.9	376	1	US-08-303-238-1	Sequence 1, Appli
53	39.5	26.9	449	2	US-08-839-008-2	Sequence 2, Appli
54	39.5	26.9	449	2	US-08-839-008-9	Sequence 9, Appli
55	39.5	26.9	1833	4	PCT-US95-02251-18	Sequence 18, Appli
56	39	26.5	179	2	US-08-606-143-33	Sequence 33, Appli
57	39	26.5	297	1	US-07-866-560-4	Sequence 4, Appli
58	39	26.5	297	1	US-08-077-673-4	Sequence 4, Appli
59	39	26.5	297	1	US-08-671-525B-4	Sequence 4, Appli
60	39	26.5	297	1	US-08-672-109B-4	Sequence 4, Appli
61	39	26.5	297	2	US-08-842-045-4	Sequence 4, Appli
62	39	26.5	297	2	US-08-842-045-4	Sequence 4, Appli
63	39	26.5	297	2	US-08-466-906B-7	Sequence 7, Appli
64	39	26.5	297	2	US-08-842-238-4	Sequence 4, Appli
65	39	26.5	297	2	US-08-780-749A-3	Sequence 3, Appli
66	39	26.5	406	3	US-08-582-740-68	Sequence 68, Appli
67	39	26.5	515	1	US-08-194-338-3	Sequence 3, Appli
68	39	26.5	515	1	US-08-444-734A-7	Sequence 7, Appli
69	39	26.5	515	2	US-08-722-001-25	Sequence 25, Appli
70	39	26.5	515	2	US-08-406-855A-22	Sequence 22, Appli
71	39	26.5	517	2	US-08-467-568-10	Sequence 10, Appli
72	39	26.5	517	2	US-09-030-582-10	Sequence 10, Appli
73	39	26.5	520	1	US-08-334-698-4	Sequence 4, Appli
74	39	26.5	520	1	US-08-228-932-4	Sequence 4, Appli
75	39	26.5	520	1	US-08-468-939-4	Sequence 4, Appli
76	39	26.5	520	2	US-08-406-855A-4	Sequence 4, Appli
77	39	26.5	520	3	US-08-722-190-4	Sequence 4, Appli
78	39	26.5	520	3	US-08-244-354-4	Sequence 4, Appli
79	39	26.5	520	4	PCT-US95-04203-4	Sequence 4, Appli
80	39	26.5	587	1	US-08-398-008A-2	Sequence 2, Appli
81	39	26.5	587	2	US-08-893-333-2	Sequence 2, Appli
82	39	26.5	793	2	US-08-720-484A-2	Sequence 2, Appli
83	39	26.5	847	1	US-08-276-099A-2	Sequence 2, Appli
84	39	26.5	847	1	US-08-781-890-2	Sequence 2, Appli
85	39	26.5	1435	2	US-08-568-459A-4	Sequence 4, Appli
86	39	26.5	1435	2	US-08-487-826B-4	Sequence 4, Appli
87	38.5	26.2	437	2	US-08-811-949-57	Sequence 57, Appli
88	38.5	26.2	682	1	US-07-998-003A-107	Sequence 107, Appli
89	38.5	26.2	682	1	US-08-453-274B-107	Sequence 107, Appli
90	38.5	26.2	682	1	US-08-453-695A-107	Sequence 107, Appli
91	38.5	26.2	682	2	US-08-268-161A-107	Sequence 107, Appli
92	38.5	26.2	682	2	US-08-453-702A-107	Sequence 107, Appli
93	38.5	26.2	682	4	PCT-US93-12588-107	Sequence 107, Appli
94	38.5	26.2	682	4	PCT-US95-08071-107	Sequence 107, Appli
95	38.5	26.2	836	1	US-07-998-003A-105	Sequence 105, Appli
96	38.5	26.2	836	1	US-08-453-274B-105	Sequence 105, Appli
97	38.5	26.2	836	1	US-08-453-695A-105	Sequence 105, Appli
98	38.5	26.2	836	1	US-08-268-161A-105	Sequence 105, Appli
99	38.5	26.2	836	2	US-08-453-702A-105	Sequence 105, Appli
100	38.5	26.2	836	4	PCT-US93-12588-105	Sequence 105, Appli
101	38.5	26.2	836	4	PCT-US95-08071-105	Sequence 105, Appli

102	38.5	26.2	904	1	US-07-998-003A-97	Sequence 97, Appl	175	37.5	25.5	527	1	US-07-609-510B-16	Sequence 16, Appl
103	38.5	26.2	904	1	US-08-453-274B-97	Sequence 97, Appl	176	37.5	25.5	527	2	US-08-811-949-39	Sequence 39, Appl
104	38.5	26.2	904	1	US-08-453-695A-97	Sequence 97, Appl	177	37.5	25.5	527	4	PCT-US91-01025A-2	Sequence 2, Appl
105	38.5	26.2	904	1	US-08-288-161A-97	Sequence 97, Appl	178	37.5	25.5	527	5	5185259-8	Patent No. 5185259
106	38.5	26.2	904	2	US-08-453-702A-97	Sequence 97, Appl	179	37.5	25.5	527	5	5520913-1	Patent No. 5520913
107	38.5	26.2	904	4	PCT-US93-12588-97	Sequence 97, Appl	180	37.5	25.5	562	2	US-08-811-949-43	Sequence 43, Appl
108	38.5	26.2	904	4	PCT-US95-08071-97	Sequence 97, Appl	181	37.5	25.5	562	2	US-08-560-098A-50	Sequence 50, Appl
109	38	25.9	20	1	US-08-182-619-11	Sequence 11, Appl	182	37.5	25.5	562	2	US-08-883-795A-38	Sequence 38, Appl
110	38	25.9	20	1	US-08-330-535A-11	Sequence 11, Appl	183	37.5	25.5	562	5	5185259-3	Patent No. 5185259
111	38	25.9	20	1	US-08-688-145-6	Sequence 6, Appl	184	37.5	25.5	562	5	5200340-2	Patent No. 5200340
112	38	25.9	20	2	US-08-838-844-11	Sequence 11, Appl	185	37.5	25.5	562	5	5244676-5	Patent No. 5244676
113	38	25.9	24	2	US-08-737-085A-14	Sequence 14, Appl	186	37.5	25.5	562	5	5344773-2	Patent No. 5344773
114	38	25.9	24	3	US-08-246-258-14	Sequence 3, Appl	187	37.5	25.5	737	1	US-08-185-432-2	Sequence 2, Appl
115	38	25.9	25	2	US-08-338-282-3	Sequence 3, Appl	188	37.5	25.5	737	1	US-08-698-551-16	Sequence 16, Appl
116	38	25.9	45	2	US-08-511-485-36	Sequence 36, Appl	189	37.5	25.5	1588	2	US-08-602-228-16	Sequence 16, Appl
117	38	25.9	342	1	US-08-499-215-2	Sequence 2, Appl	190	37.5	25.5	1588	2	US-08-839-732A-16	Sequence 16, Appl
118	38	25.9	426	2	US-08-416-870C-10	Sequence 10, Appl	191	37.5	25.5	31	1	US-08-197-792-5	Sequence 5, Appl
119	38	25.9	477	1	US-08-136-922-2	Sequence 2, Appl	192	37	25.2	31	1	US-08-459-850-5	Sequence 5, Appl
120	38	25.9	498	2	US-08-511-485-13	Sequence 13, Appl	193	37	25.2	31	1	US-08-459-214-5	Sequence 5, Appl
121	38	25.9	546	1	US-08-494-168-10	Sequence 10, Appl	194	37	25.2	31	1	US-08-197-792-11	Sequence 11, Appl
122	38	25.9	596	4	PCT-US93-00869-8	Sequence 8, Appl	195	37	25.2	60	1	US-08-459-850-11	Sequence 11, Appl
123	38	25.9	771	1	US-08-121-713D-54	Sequence 54, Appl	196	37	25.2	60	1	US-08-459-850-11	Sequence 11, Appl
124	38	25.9	771	1	US-08-835-268-54	Sequence 54, Appl	197	37	25.2	60	1	US-08-459-850-11	Sequence 11, Appl
125	38	25.9	771	2	US-08-060-692-54	Sequence 54, Appl	198	37	25.2	91	2	US-08-637-759B-60	Sequence 60, Appl
126	38	25.9	771	3	US-08-833-391-54	Sequence 54, Appl	199	37	25.2	91	3	US-08-871-355A-60	Sequence 60, Appl
127	38	25.9	771	4	PCT-US94-10151A-54	Sequence 54, Appl	200	37	25.2	121	1	US-08-047-033-1	Sequence 1, Appl
128	38	25.9	871	2	US-08-775-009-35	Sequence 35, Appl	201	37	25.2	269	1	US-08-241-766-4	Sequence 4, Appl
129	38	25.9	2050	2	US-08-347-594A-2	Sequence 2, Appl	202	37	25.2	269	1	US-08-241-766-5	Sequence 5, Appl
130	38	25.9	2476	2	US-08-276-967-2	Sequence 2, Appl	203	37	25.2	269	1	US-08-241-766-14	Sequence 14, Appl
131	38	25.9	2556	1	US-08-185-432-17	Sequence 17, Appl	204	37	25.2	289	2	US-08-741-437-4	Sequence 4, Appl
132	38	25.9	2556	1	US-08-083-590A-20	Sequence 20, Appl	205	37	25.2	289	2	US-09-134-593-4	Sequence 4, Appl
133	38	25.9	3033	1	US-07-925-695-5	Sequence 5, Appl	206	37	25.2	296	1	US-08-241-465B-21	Sequence 21, Appl
134	38	25.9	3033	1	US-07-925-695-8	Sequence 8, Appl	207	37	25.2	297	1	US-07-866-560-6	Sequence 6, Appl
135	38	25.9	3033	1	US-07-925-695-9	Sequence 9, Appl	208	37	25.2	297	1	US-08-077-673-6	Sequence 6, Appl
136	37.5	25.5	157	2	US-08-811-949-41	Sequence 41, Appl	209	37	25.2	297	1	US-08-478-992-6	Sequence 6, Appl
137	37.5	25.5	180	3	US-09-187-331-5	Sequence 5, Appl	210	37	25.2	334	1	US-08-241-465B-19	Sequence 19, Appl
138	37.5	25.5	195	3	US-09-187-331-1	Sequence 1, Appl	211	37	25.2	334	1	US-08-241-465B-20	Sequence 20, Appl
139	37.5	25.5	224	1	US-08-185-432-7	Sequence 7, Appl	212	37	25.2	334	2	US-08-484-397A-8	Sequence 8, Appl
140	37.5	25.5	285	1	US-08-292-045-5	Sequence 5, Appl	213	37	25.2	348	1	US-08-360-953A-45	Sequence 45, Appl
141	37.5	25.5	285	1	US-08-292-045-7	Sequence 7, Appl	214	37	25.2	348	2	US-08-484-397A-2	Sequence 2, Appl
142	37.5	25.5	354	2	US-08-811-949-61	Sequence 61, Appl	215	37	25.2	348	2	US-08-484-397A-3	Sequence 3, Appl
143	37.5	25.5	355	1	US-08-137-116-1	Sequence 1, Appl	216	37	25.2	348	2	US-08-484-397A-4	Sequence 4, Appl
144	37.5	25.5	355	1	US-08-217-618-1	Sequence 1, Appl	217	37	25.2	348	2	US-08-484-397A-5	Sequence 5, Appl
145	37.5	25.5	355	1	US-08-427-640-2	Sequence 2, Appl	218	37	25.2	348	2	US-08-484-397A-6	Sequence 6, Appl
146	37.5	25.5	355	1	US-08-427-640-6	Sequence 6, Appl	219	37	25.2	348	2	US-08-484-397A-7	Sequence 7, Appl
147	37.5	25.5	355	1	US-08-217-617A-1	Sequence 1, Appl	220	37	25.2	348	2	US-08-484-397A-8	Sequence 8, Appl
148	37.5	25.5	355	1	US-08-217-617A-1	Sequence 1, Appl	221	37	25.2	348	2	US-08-484-397A-27	Sequence 27, Appl
149	37.5	25.5	355	2	US-08-811-949-45	Sequence 45, Appl	222	37	25.2	375	1	US-08-205-719-2	Sequence 2, Appl
150	37.5	25.5	355	2	US-08-811-949-47	Sequence 47, Appl	223	37	25.2	376	2	US-08-608-241-2	Sequence 2, Appl
151	37.5	25.5	355	2	US-08-811-949-53	Sequence 53, Appl	224	37	25.2	376	2	US-08-922-182-2	Sequence 2, Appl
152	37.5	25.5	355	5	5223256-1	Patent No. 5223256	225	37	25.2	376	2	US-08-919-953-2	Sequence 2, Appl
153	37.5	25.5	356	1	US-08-427-640-4	Sequence 4, Appl	226	37	25.2	392	3	US-08-301-162-2	Sequence 2, Appl
154	37.5	25.5	356	1	US-08-427-640-8	Sequence 8, Appl	227	37	25.2	424	1	US-08-197-792-31	Sequence 31, Appl
155	37.5	25.5	375	2	US-08-480-994-38	Sequence 38, Appl	228	37	25.2	424	1	US-08-459-850-31	Sequence 31, Appl
156	37.5	25.5	375	2	US-08-616-844-38	Sequence 38, Appl	229	37	25.2	424	1	US-08-459-214-31	Sequence 31, Appl
157	37.5	25.5	375	2	US-08-599-654-38	Sequence 38, Appl	230	37	25.2	426	1	US-08-197-792-41	Sequence 41, Appl
158	37.5	25.5	375	2	US-08-485-573-38	Sequence 38, Appl	231	37	25.2	426	1	US-08-459-850-41	Sequence 41, Appl
159	37.5	25.5	375	3	US-08-944-423A-38	Sequence 38, Appl	232	37	25.2	426	1	US-08-459-214-41	Sequence 41, Appl
160	37.5	25.5	375	3	US-08-944-423A-38	Sequence 38, Appl	233	37	25.2	426	3	US-08-301-162-18	Sequence 18, Appl
161	37.5	25.5	389	2	US-08-811-949-65	Sequence 65, Appl	234	37	25.2	543	2	US-08-469-412A-7	Sequence 7, Appl
162	37.5	25.5	437	2	US-08-811-949-49	Sequence 49, Appl	235	37	25.2	548	2	US-08-469-412A-2	Sequence 2, Appl
163	37.5	25.5	437	2	US-08-811-949-51	Sequence 51, Appl	236	37	25.2	3111	2	US-08-460-309-4	Sequence 4, Appl
164	37.5	25.5	437	2	US-08-811-949-55	Sequence 55, Appl	237	37	25.2	3111	2	US-08-125-077-4	Sequence 4, Appl
165	37.5	25.5	472	2	US-08-811-949-63	Sequence 63, Appl	238	36.5	24.8	482	1	US-08-184-327A-8	Sequence 8, Appl
166	37.5	25.5	478	1	US-08-720-899-7	Sequence 7, Appl	239	36.5	24.8	482	4	PCT-US95-00670-8	Sequence 8, Appl
167	37.5	25.5	478	1	US-08-459-610-7	Sequence 7, Appl	240	36.5	24.8	536	1	US-08-354-456A-2	Sequence 2, Appl
168	37.5	25.5	478	2	US-08-343-804-7	Sequence 7, Appl	241	36.5	24.8	536	1	US-07-999-280A-2	Sequence 2, Appl
169	37.5	25.5	478	2	US-08-339-715A-2	Sequence 2, Appl	242	36.5	24.8	536	1	US-08-426-279-2	Sequence 2, Appl
170	37.5	25.5	478	2	US-08-600-908A-10	Sequence 10, Appl	243	36.5	24.8	536	1	US-08-401-013-2	Sequence 2, Appl
171	37.5	25.5	478	3	US-08-683-838A-10	Sequence 10, Appl	244	36.5	24.8	554	1	US-08-347-254-1	Sequence 1, Appl
172	37.5	25.5	522	1	US-08-639-237-8	Sequence 2, Appl	245	36.5	24.8	554	2	US-08-464-463-1	Sequence 1, Appl
173	37.5	25.5	522	1	US-08-975-405-2	Sequence 2, Appl	246	36.5	24.8	554	4	PCT-US93-08282-2	Sequence 2, Appl
174	37.5	25.5	526	2	US-08-852-401-3	Sequence 3, Appl	247	36.5	24.8	584	2	US-08-987-466-4	Sequence 4, Appl

248	36.5	24.8	626	1	US-08-184-327A-2	Sequence 2, Appli	321	36	24.5	691	3	US-08-946-475-2	Sequence 2, Appli
249	36.5	24.8	626	4	PCT-US95-00670-2	Sequence 2, Appli	322	36	24.5	711	3	US-08-946-475-9	Sequence 9, Appli
250	36.5	24.8	633	1	US-08-250-859-17	Sequence 17, Appl	323	36	24.5	816	2	US-08-785-310A-8	Sequence 8, Appli
251	36.5	24.8	633	1	US-08-490-803-17	Sequence 17, Appl	324	36	24.5	816	2	US-08-816-933A-53	Sequence 53, Appl
252	36.5	24.8	633	2	US-08-457-254-7	Sequence 17, Appl	325	36	24.5	976	2	US-08-449-645A-18	Sequence 18, Appl
253	36.5	24.8	633	4	PCT-US94-08806-17	Sequence 17, Appl	326	36	24.5	976	2	US-08-702-367A-18	Sequence 18, Appl
254	36.5	24.8	633	4	PCT-US95-01775-17	Sequence 17, Appl	327	36	24.5	976	4	PCT-US95-04681-18	Sequence 18, Appl
255	36.5	24.8	633	4	PCT-US95-16626-7	Sequence 7, Appli	328	36	24.5	977	2	US-08-673-789-8	Sequence 8, Appli
256	36	24.5	30	5	5258287-57	Patent No. 5258287	329	36	24.5	1064	1	US-08-537-210A-3	Sequence 3, Appli
257	36	24.5	35	1	US-08-338-992B-27	Sequence 27, Appl	330	36	24.5	1142	2	US-08-993-118-7	Sequence 7, Appli
258	36	24.5	36	1	US-08-638-271A-1	Sequence 1, Appli	331	36	24.5	1142	2	US-08-845-528C-7	Sequence 7, Appli
259	36	24.5	37	1	US-08-682-485A-5	Sequence 5, Appli	332	36	24.5	1376	4	PCT-US95-10194-3	Sequence 3, Appli
260	36	24.5	37	2	US-08-933-314-5	Sequence 5, Appli	333	36	24.5	1376	4	PCT-US95-10194-3	Sequence 3, Appli
261	36	24.5	82	1	US-08-244-116B-21	Sequence 21, Appl	334	36	24.5	2329	1	US-08-755-587-16	Sequence 16, Appl
262	36	24.5	118	1	US-08-497-025-10	Sequence 10, Appl	335	36	24.5	2523	1	US-08-185-432-18	Sequence 18, Appl
263	36	24.5	179	2	US-08-606-143-34	Sequence 34, Appl	336	36	24.5	3418	2	US-08-639-501-2	Sequence 2, Appli
264	36	24.5	180	2	US-08-786-606-3	Sequence 3, Appli	337	36	24.5	3418	2	US-08-603-753D-4	Sequence 4, Appli
265	36	24.5	180	2	US-08-933-750C-48	Sequence 48, Appl	338	36	24.5	3418	3	US-09-044-946-2	Sequence 2, Appli
266	36	24.5	187	1	US-08-067-684-14	Sequence 14, Appl	339	36	24.5	3418	3	US-08-755-587-44	Sequence 44, Appl
267	36	24.5	187	1	US-08-008-898-14	Sequence 14, Appl	340	36	24.5	4472	2	US-08-804-227C-2	Sequence 2, Appli
268	36	24.5	187	2	US-08-459-818-14	Sequence 14, Appl	341	35.5	24.1	31	1	US-08-471-780C-43	Sequence 43, Appl
269	36	24.5	187	2	US-08-889-666-14	Sequence 14, Appl	342	35.5	24.1	31	1	US-08-467-282B-43	Sequence 43, Appl
270	36	24.5	187	2	US-08-889-666-14	Sequence 14, Appl	343	35.5	24.1	31	2	US-08-471-282A-43	Sequence 43, Appl
271	36	24.5	187	2	US-08-465-078-14	Sequence 14, Appl	344	35.5	24.1	31	2	US-08-466-710C-43	Sequence 43, Appl
272	36	24.5	187	2	US-08-725-776-14	Sequence 14, Appl	345	35.5	24.1	31	3	US-08-468-739C-43	Sequence 43, Appl
273	36	24.5	187	4	PCT-US95-06726-36	Sequence 36, Appl	346	35.5	24.1	37	1	US-08-471-780C-83	Sequence 83, Appl
274	36	24.5	218	2	US-08-343-101A-20	Sequence 20, Appl	347	35.5	24.1	37	1	US-08-471-780C-91	Sequence 91, Appl
275	36	24.5	219	1	US-08-463-115-91	Sequence 91, Appl	348	35.5	24.1	37	1	US-08-467-282B-83	Sequence 83, Appl
276	36	24.5	219	1	US-08-465-388-91	Sequence 91, Appl	349	35.5	24.1	37	1	US-08-467-282B-91	Sequence 91, Appl
277	36	24.5	232	1	US-08-278-091-8	Sequence 8, Appli	350	35.5	24.1	37	2	US-08-471-282A-83	Sequence 83, Appl
278	36	24.5	232	1	US-08-483-859-8	Sequence 8, Appli	351	35.5	24.1	37	2	US-08-471-282A-91	Sequence 91, Appl
279	36	24.5	232	1	US-08-472-173-8	Sequence 8, Appli	352	35.5	24.1	37	2	US-08-466-710C-83	Sequence 83, Appl
280	36	24.5	232	2	US-08-487-167-8	Sequence 8, Appli	353	35.5	24.1	37	2	US-08-466-710C-91	Sequence 91, Appl
281	36	24.5	232	2	US-08-482-816-8	Sequence 8, Appli	354	35.5	24.1	37	3	US-08-468-739C-83	Sequence 83, Appl
282	36	24.5	232	2	US-08-296-149-8	Sequence 8, Appli	355	35.5	24.1	37	3	US-08-468-739C-91	Sequence 91, Appl
283	36	24.5	232	2	US-08-801-499-8	Sequence 8, Appli	356	35.5	24.1	58	1	US-08-334-773A-1	Sequence 1, Appli
284	36	24.5	232	2	US-08-978-404B-45	Sequence 45, Appl	357	35.5	24.1	264	1	US-08-463-115-93	Sequence 93, Appl
285	36	24.5	232	2	US-08-615-271-8	Sequence 8, Appli	358	35.5	24.1	264	1	US-08-465-388-93	Sequence 93, Appl
286	36	24.5	232	3	US-09-074-660-8	Sequence 8, Appli	359	35.5	24.1	312	1	US-08-425-061-18	Sequence 18, Appl
287	36	24.5	232	3	US-09-074-659-8	Sequence 8, Appli	360	35.5	24.1	312	2	US-08-825-886-18	Sequence 18, Appl
288	36	24.5	264	3	US-09-053-197A-8	Sequence 8, Appli	361	35.5	24.1	375	1	US-08-303-238-2	Sequence 2, Appli
289	36	24.5	267	2	US-07-857-224B-42	Sequence 42, Appl	362	35.5	24.1	516	2	US-08-794-216-3	Sequence 3, Appli
290	36	24.5	309	1	US-08-236-918A-2	Sequence 2, Appli	363	35.5	24.1	516	2	US-08-749-289-3	Sequence 3, Appli
291	36	24.5	315	1	US-08-571-758-12	Sequence 12, Appl	364	35.5	24.1	765	1	US-08-425-061-19	Sequence 19, Appl
292	36	24.5	315	1	US-08-909-984A-12	Sequence 12, Appl	365	35.5	24.1	765	1	US-08-825-886-19	Sequence 19, Appl
293	36	24.5	315	1	US-08-909-983-12	Sequence 12, Appl	366	35.5	24.1	788	1	US-08-572-225-1	Sequence 1, Appli
294	36	24.5	334	2	US-08-646-981-16	Sequence 16, Appl	367	35.5	24.1	900	1	US-08-425-061-20	Sequence 20, Appl
295	36	24.5	341	2	US-08-530-566-9	Sequence 9, Appli	368	35.5	24.1	900	2	US-08-825-886-20	Sequence 20, Appl
296	36	24.5	346	1	US-08-276-151-5	Sequence 5, Appli	369	35.5	24.1	914	1	US-08-425-061-21	Sequence 21, Appl
297	36	24.5	365	1	US-08-176-620A-2	Sequence 2, Appli	370	35.5	24.1	914	2	US-08-825-886-21	Sequence 21, Appl
298	36	24.5	365	1	US-08-463-862-2	Sequence 2, Appli	371	35.5	24.1	1093	4	PCT-US93-03077-1	Sequence 1, Appli
299	36	24.5	365	2	US-08-461-985-2	Sequence 2, Appli	372	35.5	24.1	1171	1	US-08-445-135-1	Sequence 1, Appli
300	36	24.5	365	2	US-08-458-887-2	Sequence 2, Appli	373	35.5	24.1	1202	1	US-08-425-061-22	Sequence 22, Appl
301	36	24.5	367	1	US-08-176-620A-11	Sequence 11, Appl	374	35.5	24.1	1202	2	US-08-825-886-22	Sequence 22, Appl
302	36	24.5	367	2	US-08-461-985-11	Sequence 11, Appl	375	35.5	24.1	1243	2	US-08-557-139-2	Sequence 2, Appli
303	36	24.5	382	3	US-08-582-740-70	Sequence 70, Appl	376	35.5	24.1	1363	1	US-08-425-061-23	Sequence 23, Appl
304	36	24.5	404	2	US-08-666-367B-7	Sequence 7, Appli	377	35.5	24.1	1363	2	US-08-825-886-23	Sequence 23, Appl
305	36	24.5	418	2	US-08-257-963B-2	Sequence 2, Appli	378	35.5	24.1	1852	1	US-08-425-061-24	Sequence 24, Appl
306	36	24.5	418	4	PCT-US95-07201-2	Sequence 2, Appli	379	35.5	24.1	1852	2	US-08-825-886-24	Sequence 24, Appl
307	36	24.5	469	2	US-08-484-126-1	Sequence 1, Appli	380	35.5	24.1	1863	1	US-08-558-591-2	Sequence 2, Appli
308	36	24.5	472	4	PCT-US94-09235-2	Sequence 2, Appli	381	35.5	24.1	1863	1	US-08-480-784-2	Sequence 2, Appli
309	36	24.5	490	3	US-09-039-555B-14	Sequence 14, Appl	382	35.5	24.1	1863	1	US-08-483-553-2	Sequence 2, Appli
310	36	24.5	534	3	US-08-755-587-14	Sequence 14, Appl	383	35.5	24.1	1863	1	US-08-487-002-2	Sequence 2, Appli
311	36	24.5	596	2	US-08-821-355A-6	Sequence 6, Appli	384	35.5	24.1	1863	1	US-08-483-554B-2	Sequence 2, Appli
312	36	24.5	596	2	US-09-003-687A-6	Sequence 6, Appli	385	35.5	24.1	1863	1	US-08-798-691-2	Sequence 2, Appli
313	36	24.5	648	1	US-08-276-151-2	Sequence 2, Appli	386	35.5	24.1	1863	1	US-08-798-691-4	Sequence 4, Appli
314	36	24.5	648	1	US-08-185-282-1	Sequence 1, Appli	387	35.5	24.1	1863	1	US-08-798-691-6	Sequence 6, Appli
315	36	24.5	648	1	US-08-185-282-2	Sequence 2, Appli	388	35.5	24.1	1863	1	US-08-488-011B-2	Sequence 2, Appli
316	36	24.5	648	1	US-08-185-282-3	Sequence 3, Appli	389	35.5	24.1	1863	1	US-08-603-753D-2	Sequence 2, Appli
317	36	24.5	648	1	US-08-185-282-4	Sequence 4, Appli	390	35.5	24.1	1863	4	PCT-US95-10202-2	Sequence 2, Appli
318	36	24.5	648	1	US-08-185-282-5	Sequence 5, Appli	391	35.5	24.1	1863	4	PCT-US95-10203-2	Sequence 2, Appli
319	36	24.5	648	1	US-08-185-282-12	Sequence 12, Appl	392	35.5	24.1	1863	4	PCT-US95-10220-2	Sequence 2, Appli
320	36	24.5	648	2	US-08-886-751A-6	Sequence 6, Appli	393	35	23.8	25	1	US-08-684-862-7	Sequence 7, Appli

394	35	23.8	31	1	US-08-055-917-1	Sequence 1, Appl	467	35	23.8	1251	4	PCT-US95-02251-3	Sequence 3, Appl
395	35	23.8	31	1	US-08-095-068-1	Sequence 1, Appl	468	35	23.8	1252	1	US-08-199-780-3	Sequence 3, Appl
396	35	23.8	31	1	US-07-785-565A-1	Sequence 1, Appl	469	35	23.8	1252	1	US-08-316-650-3	Sequence 3, Appl
397	35	23.8	37	1	US-08-682-485A-4	Sequence 4, Appl	470	35	23.8	1642	1	US-08-447-411-45	Sequence 45, Appl
398	35	23.8	37	2	US-08-933-314-4	Sequence 4, Appl	471	35	23.8	1642	2	US-08-662-227-2	Sequence 2, Appl
399	35	23.8	38	1	US-08-176-500-59	Sequence 59, Appl	472	35	23.8	1648	2	US-08-662-227-35	Sequence 35, Appl
400	35	23.8	38	1	US-08-471-052A-59	Sequence 59, Appl	473	35	23.8	2813	3	US-08-896-449A-2	Sequence 2, Appl
401	35	23.8	38	1	US-08-189-331-59	Sequence 59, Appl	474	34.5	23.5	58	1	US-08-334-773A-2	Sequence 2, Appl
402	35	23.8	38	2	US-08-471-939-59	Sequence 59, Appl	475	34.5	23.5	58	1	US-08-676-125A-38	Sequence 38, Appl
403	35	23.8	38	2	US-08-471-800-59	Sequence 59, Appl	476	34.5	23.5	58	1	US-09-136-012A-38	Sequence 38, Appl
404	35	23.8	38	2	US-08-471-068-59	Sequence 59, Appl	477	34.5	23.5	58	3	US-08-676-124-69	Sequence 69, Appl
405	35	23.8	120	2	US-08-869-733-4	Sequence 4, Appl	478	34.5	23.5	63	2	US-08-369-829A-18	Sequence 18, Appl
406	35	23.8	145	2	US-08-708-541A-28	Sequence 28, Appl	479	34.5	23.5	192	1	US-07-949-812-3	Sequence 3, Appl
407	35	23.8	145	2	US-08-640-977-6	Sequence 6, Appl	480	34.5	23.5	221	1	US-07-949-812-2	Sequence 2, Appl
408	35	23.8	191	2	US-08-290-665A-194	Sequence 194, App	481	34.5	23.5	323	2	US-08-747-788-2	Sequence 2, Appl
409	35	23.8	191	4	PCT-US95-10398-194	Sequence 194, App	482	34.5	23.5	323	1	US-08-457-997B-2	Sequence 2, Appl
410	35	23.8	207	2	US-08-640-977-5	Sequence 5, Appl	483	34.5	23.5	366	5	5352575-5	Patent No. 5352575
411	35	23.8	219	1	US-08-843-993-4	Sequence 4, Appl	484	34.5	23.5	402	5	US-08-331-394-2	Sequence 2, Appl
412	35	23.8	219	3	US-09-059-520A-4	Sequence 4, Appl	485	34.5	23.5	409	1	US-08-250-858-2	Sequence 2, Appl
413	35	23.8	219	3	US-09-334-275-4	Sequence 4, Appl	486	34.5	23.5	409	1	US-08-446-915-2	Sequence 2, Appl
414	35	23.8	225	2	US-08-440-517A-5	Sequence 5, Appl	487	34.5	23.5	409	1	US-08-446-915-2	Sequence 2, Appl
415	35	23.8	232	2	US-08-738-413B-11	Sequence 11, Appl	488	34.5	23.5	409	2	US-08-744-139-2	Sequence 2, Appl
416	35	23.8	248	2	US-08-640-977-2	Sequence 2, Appl	489	34.5	23.5	409	4	PCT-US95-06639-2	Sequence 2, Appl
417	35	23.8	266	2	US-08-640-977-4	Sequence 4, Appl	490	34.5	23.5	557	1	US-08-309-341-2	Sequence 2, Appl
418	35	23.8	326	2	US-08-640-977-1	Sequence 1, Appl	491	34.5	23.5	557	1	US-08-309-341-4	Sequence 4, Appl
419	35	23.8	343	2	US-08-788-539A-2	Sequence 2, Appl	492	34.5	23.5	557	1	US-08-608-267-2	Sequence 2, Appl
420	35	23.8	362	1	US-08-247-907A-11	Sequence 11, Appl	493	34.5	23.5	557	1	US-08-608-267-4	Sequence 2, Appl
421	35	23.8	362	1	US-08-452-772-11	Sequence 11, Appl	494	34.5	23.5	557	1	US-08-608-452-2	Sequence 2, Appl
422	35	23.8	362	4	PCT-US94-05288-11	Sequence 11, Appl	495	34.5	23.5	557	1	US-08-608-452-4	Sequence 4, Appl
423	35	23.8	387	1	US-08-123-161A-10	Sequence 10, Appl	496	34.5	23.5	557	1	US-08-608-224-2	Sequence 2, Appl
424	35	23.8	387	1	US-08-483-278-10	Sequence 10, Appl	497	34.5	23.5	557	1	US-08-608-224-4	Sequence 4, Appl
425	35	23.8	407	2	US-08-765-875-2	Sequence 2, Appl	498	34.5	23.5	557	2	US-08-967-149-2	Sequence 2, Appl
426	35	23.8	407	2	US-08-765-875-6	Sequence 6, Appl	499	34.5	23.5	557	2	US-08-967-149-4	Sequence 4, Appl
427	35	23.8	407	3	US-08-795-671-2	Sequence 2, Appl	500	34.5	23.5	599	2	US-08-426-125-1	Sequence 1, Appl
428	35	23.8	407	3	US-08-795-671-6	Sequence 6, Appl	501	34.5	23.5	599	2	US-08-426-125-3	Sequence 3, Appl
429	35	23.8	410	1	US-08-073-807A-16	Sequence 16, Appl	502	34.5	23.5	599	2	US-08-455-355-1	Sequence 1, Appl
430	35	23.8	422	2	US-09-096-982-5	Sequence 5, Appl	503	34.5	23.5	599	2	US-08-455-355-3	Sequence 3, Appl
431	35	23.8	422	2	US-08-653-650A-5	Sequence 5, Appl	504	34.5	23.5	638	2	US-08-426-125-4	Sequence 4, Appl
432	35	23.8	434	1	US-08-111-939-13	Sequence 13, Appl	505	34.5	23.5	638	2	US-08-426-125-5	Sequence 5, Appl
433	35	23.8	435	1	US-08-111-939-14	Sequence 14, Appl	506	34.5	23.5	638	2	US-08-455-355-4	Sequence 4, Appl
434	35	23.8	435	1	US-08-111-939-15	Sequence 15, Appl	507	34.5	23.5	638	2	US-08-455-355-5	Sequence 5, Appl
435	35	23.8	435	1	US-08-111-939-16	Sequence 16, Appl	508	34.5	23.5	1135	2	US-08-469-537A-97	Sequence 97, Appl
436	35	23.8	435	1	US-08-452-262-2	Sequence 2, Appl	509	34.5	23.5	1390	2	US-08-770-544-2	Sequence 2, Appl
437	35	23.8	435	1	US-08-734-550-2	Sequence 2, Appl	510	34.5	23.5	3174	2	US-08-477-451-3	Sequence 3, Appl
438	35	23.8	435	4	PCT-US96-07558-2	Sequence 2, Appl	511	34	23.1	25	1	US-08-264-030-8	Sequence 8, Appl
439	35	23.8	443	2	US-09-096-982-9	Sequence 9, Appl	512	34	23.1	40	1	US-08-144-121-11	Sequence 11, Appl
440	35	23.8	443	2	US-08-653-650A-9	Sequence 9, Appl	513	34	23.1	40	2	US-08-735-893-11	Sequence 11, Appl
441	35	23.8	453	2	US-08-599-171A-27	Sequence 27, Appl	514	34	23.1	45	3	US-09-040-285A-9	Sequence 9, Appl
442	35	23.8	453	2	US-08-646-590B-27	Sequence 27, Appl	515	34	23.1	46	2	US-08-511-485-32	Sequence 32, Appl
443	35	23.8	453	3	US-09-069-226-27	Sequence 27, Appl	516	34	23.1	46	2	US-08-511-485-33	Sequence 33, Appl
444	35	23.8	469	2	US-08-416-870C-6	Sequence 6, Appl	517	34	23.1	52	4	PCT-US96-12860-11	Sequence 11, Appl
445	35	23.8	474	2	US-09-096-982-8	Sequence 8, Appl	518	34	23.1	52	4	PCT-US96-12860-12	Sequence 12, Appl
446	35	23.8	474	2	US-08-653-650A-8	Sequence 8, Appl	519	34	23.1	80	3	US-09-040-285A-3	Sequence 3, Appl
447	35	23.8	489	2	US-08-794-795-7	Sequence 7, Appl	520	34	23.1	96	3	US-08-813-884-39	Sequence 39, Appl
448	35	23.8	513	2	US-08-406-855A-21	Sequence 21, Appl	521	34	23.1	105	2	US-08-826-910-3	Sequence 3, Appl
449	35	23.8	518	1	US-08-392-367B-2	Sequence 2, Appl	522	34	23.1	124	1	US-08-244-116B-15	Sequence 15, Appl
450	35	23.8	519	2	US-08-725-736D-2	Sequence 2, Appl	523	34	23.1	131	2	US-08-650-598-3	Sequence 3, Appl
451	35	23.8	601	2	US-08-795-868-16	Sequence 16, Appl	524	34	23.1	144	1	US-08-133-979A-21	Sequence 21, Appl
452	35	23.8	661	2	US-08-795-868-14	Sequence 14, Appl	525	34	23.1	144	2	US-08-436-890-21	Sequence 21, Appl
453	35	23.8	766	3	US-08-539-205A-4	Sequence 4, Appl	526	34	23.1	144	2	US-08-451-213-21	Sequence 21, Appl
454	35	23.8	769	2	US-08-789-078-1	Sequence 1, Appl	527	34	23.1	149	1	US-08-425-673-5	Sequence 5, Appl
455	35	23.8	769	2	US-08-752-633-1	Sequence 1, Appl	528	34	23.1	175	2	US-08-970-133-5	Sequence 3, Appl
456	35	23.8	769	2	US-08-476-062A-45	Sequence 45, Appl	529	34	23.1	179	2	US-08-606-143-35	Sequence 35, Appl
457	35	23.8	769	2	US-07-728-215-31	Sequence 31, Appl	530	34	23.1	179	2	US-08-606-143-36	Sequence 36, Appl
458	35	23.8	769	4	PCT-US95-04886-1	Sequence 1, Appl	531	34	23.1	191	2	US-08-290-665A-187	Sequence 187, App
459	35	23.8	855	2	PCT-US96-01314-45	Sequence 45, Appl	532	34	23.1	191	2	US-08-290-665A-188	Sequence 188, App
460	35	23.8	867	2	US-08-938-365-3	Sequence 3, Appl	533	34	23.1	191	2	US-08-290-665A-189	Sequence 189, App
461	35	23.8	867	3	US-08-938-365-2	Sequence 2, Appl	534	34	23.1	191	2	US-08-290-665A-190	Sequence 190, App
462	35	23.8	885	3	US-09-074-579-5	Sequence 5, Appl	535	34	23.1	191	2	US-08-290-665A-192	Sequence 192, App
463	35	23.8	954	2	US-08-749-169A-3	Sequence 3, Appl	536	34	23.1	191	2	US-08-290-665A-193	Sequence 193, App
464	35	23.8	954	2	US-09-130-032A-3	Sequence 3, Appl	537	34	23.1	191	2	US-08-290-665A-195	Sequence 195, App
465	35	23.8	1170	1	US-08-313-288B-20	Sequence 20, Appl	538	34	23.1	191	2	US-08-290-665A-196	Sequence 196, App
466	35	23.8	1218	2	US-08-400-159-6	Sequence 6, Appl	539	34	23.1	191	4	PCT-US95-10398-187	Sequence 187, App

540	34	23.1	191	4	PCT-US95-10398-188	Sequence 188, App	613	34	23.1	618	4	PCT-US96-12860-2	Sequence 2, Appli
541	34	23.1	191	4	PCT-US95-10398-189	Sequence 189, App	614	34	23.1	648	1	US-08-451-715A-4	Sequence 4, Appli
542	34	23.1	191	4	PCT-US95-10398-190	Sequence 190, App	615	34	23.1	650	1	US-08-121-7130-60	Sequence 60, Appl
543	34	23.1	191	4	PCT-US95-10398-192	Sequence 192, App	616	34	23.1	650	1	US-08-835-268-60	Sequence 60, Appl
544	34	23.1	191	4	PCT-US95-10398-193	Sequence 193, App	617	34	23.1	650	3	US-09-060-692-60	Sequence 60, Appl
545	34	23.1	191	4	PCT-US95-10398-195	Sequence 195, App	618	34	23.1	650	3	US-08-833-391-60	Sequence 60, Appl
546	34	23.1	191	4	PCT-US95-10398-196	Sequence 196, App	619	34	23.1	650	4	PCT-US94-10151A-60	Sequence 60, Appl
547	34	23.1	209	3	US-09-040-285A-8	Sequence 8, Appli	620	34	23.1	667	2	US-08-718-661-2	Sequence 2, Appli
548	34	23.1	217	2	US-08-622-720A-24	Sequence 24, Appl	621	34	23.1	686	3	US-09-306-922-2	Sequence 2, Appli
549	34	23.1	218	1	US-08-463-115-92	Sequence 92, Appl	622	34	23.1	695	2	US-08-701-240-4	Sequence 4, Appli
550	34	23.1	218	1	US-08-465-388-92	Sequence 92, Appl	623	34	23.1	716	2	US-08-766-982-1	Sequence 1, Appli
551	34	23.1	222	1	US-08-328-152A-8	Sequence 8, Appli	624	34	23.1	765	1	US-08-309-512-5	Sequence 5, Appli
552	34	23.1	229	3	US-09-040-285A-7	Sequence 7, Appli	625	34	23.1	765	4	PCT-US92-08756A-5	Sequence 5, Appli
553	34	23.1	236	2	US-08-738-413B-10	Sequence 10, Appl	626	34	23.1	798	2	US-07-728-215-30	Sequence 30, Appl
554	34	23.1	239	4	PCT-US93-01652-1	Sequence 1, Appli	627	34	23.1	846	1	US-08-149-103-3	Sequence 3, Appli
555	34	23.1	240	1	US-07-940-861-12	Sequence 12, Appl	628	34	23.1	846	1	US-08-149-103-4	Sequence 4, Appli
556	34	23.1	240	1	US-08-459-512-12	Sequence 12, Appl	629	34	23.1	846	1	US-08-451-883-3	Sequence 3, Appli
557	34	23.1	240	1	US-08-459-657-12	Sequence 12, Appl	630	34	23.1	846	1	US-08-451-883-4	Sequence 4, Appli
558	34	23.1	240	2	US-08-460-132-12	Sequence 12, Appl	631	34	23.1	873	1	US-08-393-734-2	Sequence 2, Appli
559	34	23.1	240	4	PCT-US92-02050-12	Sequence 12, Appl	632	34	23.1	926	3	US-08-755-587-187	Sequence 187, App
560	34	23.1	240	5	5185441-36	Sequence 1, Appli	633	34	23.1	973	2	US-08-683-262B-75	Sequence 75, Appl
561	34	23.1	240	5	5223394-4	Patent No. 5185441	634	34	23.1	1111	1	US-08-317-450B-15	Sequence 15, Appl
562	34	23.1	240	5	5223394-6	Patent No. 5223394	635	34	23.1	1144	1	US-08-147-812-5	Sequence 5, Appli
563	34	23.1	244	1	US-08-910-973-22	Sequence 22, Appl	636	34	23.1	1144	2	US-08-319-866-12	Sequence 12, Appl
564	34	23.1	250	1	US-07-940-861-10	Sequence 10, Appl	637	34	23.1	1193	1	US-08-317-450B-13	Sequence 13, Appl
565	34	23.1	250	1	US-08-459-512-10	Sequence 10, Appl	638	34	23.1	1237	2	US-08-241-853-2	Sequence 2, Appli
566	34	23.1	250	2	US-08-459-657-10	Sequence 10, Appl	639	34	23.1	1237	2	US-08-850-917-2	Sequence 2, Appli
567	34	23.1	250	2	US-08-460-132-10	Sequence 10, Appl	640	34	23.1	1377	2	US-08-308-818-4	Sequence 4, Appli
568	34	23.1	250	4	PCT-US92-02050-10	Sequence 10, Appl	641	34	23.1	1404	2	US-08-400-159-2	Sequence 2, Appli
569	34	23.1	250	5	5223394-1	Patent No. 5223394	642	34	23.1	1745	2	US-09-031-485-33	Sequence 33, Appl
570	34	23.1	251	5	5185441-38	Patent No. 5185441	643	34	23.1	1745	1	US-08-847-429A-33	Sequence 33, Appl
571	34	23.1	269	1	US-08-241-766-6	Sequence 6, Appli	644	34	23.1	2414	1	US-08-227-536-2	Sequence 2, Appli
572	34	23.1	269	1	US-08-241-766-7	Sequence 7, Appli	645	34	23.1	2414	4	PCT-US95-04682-2	Sequence 2, Appli
573	34	23.1	270	2	US-09-031-485-7	Sequence 7, Appli	646	34	23.1	15281	2	US-08-471-119A-2	Sequence 2, Appli
574	34	23.1	312	2	US-08-847-429A-7	Sequence 7, Appli	647	33.5	22.8	25	2	US-08-637-759B-351	Sequence 351, App
575	34	23.1	312	2	US-09-031-485-2	Sequence 2, Appli	648	33.5	22.8	25	3	US-08-871-355A-351	Sequence 351, App
576	34	23.1	312	2	US-08-847-429A-2	Sequence 2, Appli	649	33.5	22.8	31	1	US-08-066-325-76	Sequence 76, Appl
577	34	23.1	320	2	US-07-841-591A-15	Sequence 15, Appl	650	33.5	22.8	31	1	US-08-066-325-77	Sequence 77, Appl
578	34	23.1	320	4	PCT-US93-02034-15	Sequence 15, Appl	651	33.5	22.8	31	1	US-08-066-325-78	Sequence 78, Appl
579	34	23.1	328	5	5212074-4	Patent No. 5212074	652	33.5	22.8	51	1	US-08-173-102-2	Sequence 2, Appli
580	34	23.1	341	2	US-08-530-566-7	Sequence 7, Appli	653	33.5	22.8	51	1	PCT-US94-07779-13	Sequence 13, Appl
581	34	23.1	346	4	PCT-US96-10602-2	Sequence 2, Appli	654	33.5	22.8	51	4	PCT-US94-14617-2	Sequence 2, Appli
582	34	23.1	351	2	US-08-929-417-2	Sequence 2, Appli	655	33.5	22.8	77	1	US-08-420-235B-7	Sequence 7, Appli
583	34	23.1	351	4	PCT-US91-06418-1	Sequence 1, Appli	656	33.5	22.8	77	4	PCT-US95-10194-7	Sequence 7, Appli
584	34	23.1	359	1	US-07-662-005A-4	Sequence 4, Appli	657	33.5	22.8	119	1	US-08-340-539A-18	Sequence 18, Appl
585	34	23.1	373	2	US-08-599-171A-26	Sequence 26, Appl	658	33.5	22.8	153	3	US-08-840-146-10	Sequence 10, Appl
586	34	23.1	373	2	US-08-646-590B-26	Sequence 26, Appl	659	33.5	22.8	184	1	US-08-173-102-1	Sequence 1, Appli
587	34	23.1	373	3	US-09-069-226-26	Sequence 26, Appl	660	33.5	22.8	184	4	PCT-US94-14617-1	Sequence 1, Appli
588	34	23.1	432	2	US-08-677-049-8	Sequence 8, Appli	661	33.5	22.8	192	2	US-08-531-525-29	Sequence 29, Appl
589	34	23.1	438	4	PCT-US95-05922A-2	Sequence 2, Appli	662	33.5	22.8	192	2	US-08-718-270A-29	Sequence 29, Appl
590	34	23.1	441	4	PCT-US93-12588-98	Sequence 98, Appl	663	33.5	22.8	213	3	US-08-718-738-18	Sequence 18, Appl
591	34	23.1	441	4	PCT-US95-08071-98	Sequence 98, Appl	664	33.5	22.8	330	1	US-08-468-853-6	Sequence 6, Appli
592	34	23.1	512	1	US-07-779-890-4	Sequence 4, Appli	665	33.5	22.8	330	1	US-08-468-855-6	Sequence 6, Appli
593	34	23.1	512	1	US-07-779-890-4	Sequence 4, Appli	666	33.5	22.8	330	1	US-08-310-357-6	Sequence 6, Appli
594	34	23.1	512	2	US-09-008-962-4	Sequence 4, Appli	667	33.5	22.8	330	1	US-08-468-852-6	Sequence 6, Appli
595	34	23.1	512	4	PCT-US93-05640-4	Sequence 4, Appli	668	33.5	22.8	330	2	US-08-468-857-6	Sequence 6, Appli
596	34	23.1	512	2	US-08-464-340A-13	Sequence 13, Appl	669	33.5	22.8	342	1	US-08-499-215-5	Sequence 5, Appli
597	34	23.1	539	1	US-08-484-493-2	Sequence 2, Appli	670	33.5	22.8	367	1	US-08-257-341-5	Sequence 5, Appli
598	34	23.1	550	1	US-08-484-493-2	Sequence 2, Appli	671	33.5	22.8	372	2	US-08-360-606B-31	Sequence 31, Appl
599	34	23.1	550	1	US-08-484-494-2	Sequence 2, Appli	672	33.5	22.8	469	2	US-08-968-751-2	Sequence 2, Appli
600	34	23.1	550	2	US-08-345-212-2	Sequence 2, Appli	673	33.5	22.8	506	3	US-08-840-146-18	Sequence 18, Appl
601	34	23.1	555	2	US-08-453-702A-98	Sequence 98, Appl	674	33.5	22.8	507	4	PCT-US93-08386-8	Sequence 8, Appli
602	34	23.1	556	1	US-07-998-003A-98	Sequence 98, Appl	675	33.5	22.8	514	3	US-08-840-146-2	Sequence 2, Appli
603	34	23.1	556	1	US-08-453-274B-98	Sequence 98, Appl	676	33.5	22.8	664	2	US-08-426-125-7	Sequence 7, Appli
604	34	23.1	556	1	US-08-453-695A-98	Sequence 98, Appl	677	33.5	22.8	664	2	US-08-455-355-7	Sequence 7, Appli
605	34	23.1	556	1	US-08-268-161A-98	Sequence 98, Appl	678	33.5	22.8	667	2	US-08-426-125-6	Sequence 6, Appli
606	34	23.1	564	2	US-08-865-337A-1	Sequence 1, Appli	679	33.5	22.8	667	2	US-08-455-355-6	Sequence 6, Appli
607	34	23.1	593	1	US-08-296-362-2	Sequence 2, Appli	680	33.5	22.8	670	1	US-08-363-208-2	Sequence 2, Appli
608	34	23.1	604	2	US-08-511-485-6	Sequence 6, Appli	681	33.5	22.8	671	2	US-08-426-125-8	Sequence 8, Appli
609	34	23.1	604	4	PCT-US96-12860-4	Sequence 4, Appli	682	33.5	22.8	671	2	US-08-455-355-8	Sequence 8, Appli
610	34	23.1	610	2	US-08-865-337A-6	Sequence 6, Appli	683	33.5	22.8	875	1	US-08-480-547A-23	Sequence 23, Appl
611	34	23.1	612	4	PCT-US96-12860-14	Sequence 14, Appl	684	33.5	22.8	875	1	US-08-250-847B-23	Sequence 23, Appl
612	34	23.1	618	2	US-08-511-485-8	Sequence 8, Appli	685	33.5	22.8	875	2	US-08-463-949A-23	Sequence 23, Appl

686	33.5	22.8	875	3	US-08-464-410A-23	Sequence 23, Appl	759	33	22.4	167	4	PCT-US95-07289-2	Sequence 2, Appl
687	33.5	22.8	875	4	PCT-US94-06066-23	Sequence 23, Appl	760	33	22.4	176	1	US-08-414-926A-18	Sequence 18, Appl
688	33	22.4	10	2	US-08-849-480A-19	Sequence 19, Appl	761	33	22.4	176	2	US-08-809-367-10	Sequence 10, Appl
689	33	22.4	12	1	US-08-338-992B-29	Sequence 29, Appl	762	33	22.4	176	2	US-08-926-322-18	Sequence 18, Appl
690	33	22.4	14	2	US-08-107-676-8	Sequence 8, Appl	763	33	22.4	176	3	US-09-253-682-18	Sequence 18, Appl
691	33	22.4	15	2	US-08-392-338A-5	Sequence 5, Appl	764	33	22.4	176	4	PCT-US95-13662A-10	Sequence 10, Appl
692	33	22.4	15	2	US-08-997-080-22	Sequence 22, Appl	765	33	22.4	177	2	US-08-583-562B-6	Sequence 6, Appl
693	33	22.4	15	2	US-08-997-080-24	Sequence 24, Appl	766	33	22.4	177	2	US-08-779-113-6	Sequence 6, Appl
694	33	22.4	15	2	US-08-997-362-22	Sequence 22, Appl	767	33	22.4	179	2	US-08-845-539-6	Sequence 6, Appl
695	33	22.4	15	2	US-08-997-362-24	Sequence 24, Appl	768	33	22.4	190	1	US-07-681-701-16	Sequence 16, Appl
696	33	22.4	15	3	US-08-873-970-22	Sequence 22, Appl	769	33	22.4	190	1	US-08-290-665A-155	Sequence 155, App
697	33	22.4	15	3	US-08-873-970-24	Sequence 24, Appl	770	33	22.4	191	2	US-08-290-665A-156	Sequence 156, App
698	33	22.4	15	3	US-09-166-750-5	Sequence 5, Appl	771	33	22.4	191	2	US-08-290-665A-157	Sequence 157, App
699	33	22.4	15	3	US-09-166-093-5	Sequence 5, Appl	772	33	22.4	191	2	US-08-290-665A-158	Sequence 158, App
700	33	22.4	19	2	US-08-997-080-23	Sequence 23, Appl	773	33	22.4	191	2	US-08-290-665A-159	Sequence 159, App
701	33	22.4	19	2	US-08-997-362-23	Sequence 23, Appl	774	33	22.4	191	2	US-08-290-665A-160	Sequence 160, App
702	33	22.4	19	3	US-08-873-970-23	Sequence 23, Appl	775	33	22.4	191	2	US-08-290-665A-161	Sequence 161, App
703	33	22.4	19	5	5185431-23	Patent No. 5185431	776	33	22.4	191	2	US-08-290-665A-162	Sequence 162, App
704	33	22.4	22	2	US-08-737-085A-13	Sequence 13, Appl	777	33	22.4	191	2	US-08-290-665A-163	Sequence 163, App
705	33	22.4	22	3	US-09-246-258-13	Sequence 13, Appl	778	33	22.4	191	2	US-08-290-665A-164	Sequence 164, App
706	33	22.4	26	2	US-08-620-151-85	Sequence 85, Appl	779	33	22.4	191	2	US-08-290-665A-165	Sequence 165, App
707	33	22.4	26	2	US-08-107-676-38	Sequence 38, Appl	780	33	22.4	191	2	US-08-290-665A-166	Sequence 166, App
708	33	22.4	32	2	US-08-447-430A-33	Sequence 33, Appl	781	33	22.4	191	2	US-08-290-665A-167	Sequence 167, App
709	33	22.4	37	1	US-08-290-448A-29	Sequence 29, Appl	782	33	22.4	191	2	US-08-290-665A-168	Sequence 168, App
710	33	22.4	37	1	US-08-290-448A-29	Sequence 29, Appl	783	33	22.4	191	2	US-08-290-665A-169	Sequence 169, App
711	33	22.4	37	1	US-08-175-069A-29	Sequence 29, Appl	784	33	22.4	191	2	US-08-290-665A-170	Sequence 170, App
712	33	22.4	40	4	PCT-US95-02373-15	Sequence 15, Appl	785	33	22.4	191	2	US-08-290-665A-172	Sequence 172, App
713	33	22.4	43	2	US-08-499-676A-32	Sequence 32, Appl	786	33	22.4	191	2	US-08-290-665A-173	Sequence 173, App
714	33	22.4	60	1	US-07-681-703B-22	Sequence 22, Appl	787	33	22.4	191	2	US-08-290-665A-174	Sequence 174, App
715	33	22.4	60	2	US-08-407-410B-22	Sequence 22, Appl	788	33	22.4	191	2	US-08-290-665A-175	Sequence 175, App
716	33	22.4	60	2	US-08-485-500-22	Sequence 22, Appl	789	33	22.4	191	2	US-08-290-665A-176	Sequence 176, App
717	33	22.4	60	4	PCT-US91-02370-22	Sequence 22, Appl	790	33	22.4	191	2	US-08-290-665A-176	Sequence 176, App
718	33	22.4	77	2	US-08-570-227A-6	Sequence 6, Appl	791	33	22.4	191	4	PCT-US95-10398-155	Sequence 155, App
719	33	22.4	78	1	US-08-264-534-2	Sequence 2, Appl	792	33	22.4	191	4	PCT-US95-10398-156	Sequence 156, App
720	33	22.4	78	1	US-08-083-590A-15	Sequence 15, Appl	793	33	22.4	191	4	PCT-US95-10398-157	Sequence 157, App
721	33	22.4	78	1	US-08-465-500-2	Sequence 2, Appl	794	33	22.4	191	4	PCT-US95-10398-158	Sequence 158, App
722	33	22.4	78	2	US-08-346-126-2	Sequence 2, Appl	795	33	22.4	191	4	PCT-US95-10398-159	Sequence 159, App
723	33	22.4	78	2	US-08-346-128-2	Sequence 2, Appl	796	33	22.4	191	4	PCT-US95-10398-160	Sequence 160, App
724	33	22.4	79	2	US-08-499-676A-13	Sequence 13, Appl	797	33	22.4	191	4	PCT-US95-10398-161	Sequence 161, App
725	33	22.4	82	3	US-08-474-700B-44	Sequence 44, Appl	798	33	22.4	191	4	PCT-US95-10398-162	Sequence 162, App
726	33	22.4	86	2	US-08-785-795-1	Sequence 1, Appl	799	33	22.4	191	4	PCT-US95-10398-163	Sequence 163, App
727	33	22.4	89	1	US-07-681-703B-24	Sequence 24, Appl	800	33	22.4	191	4	PCT-US95-10398-164	Sequence 164, App
728	33	22.4	89	2	US-08-407-410B-24	Sequence 24, Appl	801	33	22.4	191	4	PCT-US95-10398-165	Sequence 165, App
729	33	22.4	89	2	US-08-485-500-24	Sequence 24, Appl	802	33	22.4	191	4	PCT-US95-10398-166	Sequence 166, App
730	33	22.4	89	4	PCT-US91-02370-24	Sequence 24, Appl	803	33	22.4	191	4	PCT-US95-10398-167	Sequence 167, App
731	33	22.4	90	1	US-07-681-703B-20	Sequence 20, Appl	804	33	22.4	191	4	PCT-US95-10398-168	Sequence 168, App
732	33	22.4	90	2	US-08-407-410B-20	Sequence 20, Appl	805	33	22.4	191	4	PCT-US95-10398-169	Sequence 169, App
733	33	22.4	90	2	US-08-485-500-20	Sequence 20, Appl	806	33	22.4	191	4	PCT-US95-10398-170	Sequence 170, App
734	33	22.4	90	4	PCT-US91-02370-20	Sequence 20, Appl	807	33	22.4	191	4	PCT-US95-10398-172	Sequence 172, App
735	33	22.4	102	2	US-08-537-811-42	Sequence 42, Appl	808	33	22.4	191	4	PCT-US95-10398-173	Sequence 173, App
736	33	22.4	108	1	US-08-485-359-2	Sequence 2, Appl	809	33	22.4	191	4	PCT-US95-10398-174	Sequence 174, App
737	33	22.4	108	1	US-08-569-594-2	Sequence 2, Appl	810	33	22.4	191	4	PCT-US95-10398-175	Sequence 175, App
738	33	22.4	108	4	PCT-US96-08815-2	Sequence 4, Appl	811	33	22.4	191	4	PCT-US95-10398-176	Sequence 176, App
739	33	22.4	109	1	US-08-485-359-4	Sequence 4, Appl	812	33	22.4	191	4	PCT-US95-10398-176	Sequence 176, App
740	33	22.4	109	1	US-08-569-594-4	Sequence 4, Appl	813	33	22.4	205	5	5185431-1	Patent No. 5185431
741	33	22.4	109	4	PCT-US96-08815-4	Sequence 4, Appl	814	33	22.4	215	1	US-07-681-703B-12	Sequence 12, Appl
742	33	22.4	115	1	US-08-324-977-8	Sequence 8, Appl	815	33	22.4	215	1	US-07-681-703B-14	Sequence 14, Appl
743	33	22.4	115	2	US-08-384-616-8	Sequence 8, Appl	816	33	22.4	215	4	PCT-US91-02370-12	Sequence 12, Appl
744	33	22.4	115	2	US-08-904-686A-8	Sequence 8, Appl	817	33	22.4	215	4	PCT-US91-02370-14	Sequence 14, Appl
745	33	22.4	119	1	US-07-681-703B-18	Sequence 18, Appl	818	33	22.4	217	2	US-08-407-410B-12	Sequence 12, Appl
746	33	22.4	119	2	US-08-407-410B-18	Sequence 18, Appl	819	33	22.4	217	2	US-08-407-410B-14	Sequence 14, Appl
747	33	22.4	119	2	US-08-485-500-18	Sequence 18, Appl	820	33	22.4	217	2	US-08-485-500-12	Sequence 12, Appl
748	33	22.4	119	4	PCT-US91-02370-18	Sequence 18, Appl	821	33	22.4	217	2	US-08-485-500-14	Sequence 14, Appl
749	33	22.4	124	2	US-08-477-451-27	Sequence 27, Appl	822	33	22.4	222	5	5185431-4	Patent No. 5185431
750	33	22.4	150	1	US-07-681-703B-16	Sequence 16, Appl	823	33	22.4	223	1	US-08-143-579A-4	Sequence 4, Appl
751	33	22.4	150	2	US-08-407-410B-16	Sequence 16, Appl	824	33	22.4	223	1	US-08-143-579A-4	Sequence 4, Appl
752	33	22.4	150	2	US-08-485-500-16	Sequence 16, Appl	825	33	22.4	248	2	US-08-671-978A-9	Sequence 9, Appl
753	33	22.4	150	4	PCT-US91-02370-16	Sequence 16, Appl	826	33	22.4	259	1	US-07-795-859B-4	Sequence 4, Appl
754	33	22.4	151	2	US-08-722-050-6	Sequence 6, Appl	827	33	22.4	259	1	US-08-457-616-4	Sequence 4, Appl
755	33	22.4	155	5	5338678-1	Patent No. 5338678	828	33	22.4	266	2	US-07-857-224B-43	Sequence 43, Appl
756	33	22.4	158	5	5338678-2	Patent No. 5338678	829	33	22.4	266	2	US-07-857-224B-43	Sequence 44, Appl
757	33	22.4	165	2	US-08-602-208-3	Sequence 3, Appl	830	33	22.4	272	1	US-08-117-083-13	Sequence 13, Appl
758	33	22.4	167	1	US-08-469-667-2	Sequence 2, Appl	831	33	22.4	307	3	US-08-508-761B-33	Sequence 33, Appl

832	33	22.4	308	3	US-08-508-761B-35	Sequence 35, Appl	905	33	22.4	442	2	US-09-003-687A-5	Sequence 5, Appl
833	33	22.4	309	3	US-08-508-761B-32	Sequence 32, Appl	906	33	22.4	445	4	PCT-US94-05387-7	Sequence 7, Appl
834	33	22.4	310	3	US-08-651-136C-22	Sequence 22, Appl	907	33	22.4	448	1	US-08-207-904-2	Sequence 2, Appl
835	33	22.4	313	3	US-08-508-761B-34	Sequence 34, Appl	908	33	22.4	448	1	US-08-207-904-17	Sequence 17, Appl
836	33	22.4	321	3	US-08-508-761B-36	Sequence 36, Appl	909	33	22.4	465	3	US-08-701-582B-13	Sequence 13, Appl
837	33	22.4	321	3	US-08-508-761B-37	Sequence 37, Appl	910	33	22.4	465	3	US-09-063-893A-19	Sequence 19, Appl
838	33	22.4	323	2	US-08-997-080-35	Sequence 35, Appl	911	33	22.4	468	2	US-08-303-569B-7	Sequence 7, Appl
839	33	22.4	323	2	US-08-997-362-35	Sequence 35, Appl	912	33	22.4	468	2	US-08-116-247-7	Sequence 7, Appl
840	33	22.4	323	3	US-08-873-970-35	Sequence 35, Appl	913	33	22.4	470	2	US-08-466-906B-8	Sequence 8, Appl
841	33	22.4	325	2	US-08-107-676-29	Sequence 29, Appl	914	33	22.4	483	2	US-08-416-870C-4	Sequence 4, Appl
842	33	22.4	325	2	US-08-107-676-30	Sequence 30, Appl	915	33	22.4	484	2	US-08-252-493C-9	Sequence 9, Appl
843	33	22.4	325	2	US-08-997-080-33	Sequence 33, Appl	916	33	22.4	484	3	US-09-376-197-9	Sequence 9, Appl
844	33	22.4	325	2	US-08-997-362-33	Sequence 33, Appl	917	33	22.4	485	2	US-08-416-870C-2	Sequence 2, Appl
845	33	22.4	325	2	US-08-873-970-33	Sequence 33, Appl	918	33	22.4	486	2	US-08-821-355A-8	Sequence 8, Appl
846	33	22.4	326	2	US-08-997-080-43	Sequence 43, Appl	919	33	22.4	486	2	US-09-003-687A-8	Sequence 8, Appl
847	33	22.4	326	2	US-08-997-362-43	Sequence 43, Appl	920	33	22.4	495	2	US-08-749-902-3	Sequence 3, Appl
848	33	22.4	326	3	US-08-873-970-43	Sequence 43, Appl	921	33	22.4	506	2	US-08-820-170A-19	Sequence 19, Appl
849	33	22.4	327	2	US-08-997-080-31	Sequence 31, Appl	922	33	22.4	511	2	US-08-821-355A-9	Sequence 9, Appl
850	33	22.4	327	2	US-08-997-362-31	Sequence 31, Appl	923	33	22.4	511	2	US-09-003-687A-9	Sequence 9, Appl
851	33	22.4	327	3	US-08-873-970-31	Sequence 31, Appl	924	33	22.4	512	2	US-08-828-596-2	Sequence 2, Appl
852	33	22.4	330	2	US-08-997-080-30	Sequence 30, Appl	925	33	22.4	529	2	US-08-933-227-4	Sequence 4, Appl
853	33	22.4	330	2	US-08-997-362-30	Sequence 30, Appl	926	33	22.4	531	2	US-08-724-974A-2	Sequence 2, Appl
854	33	22.4	333	3	US-08-873-970-30	Sequence 30, Appl	927	33	22.4	533	1	US-08-368-071-10	Sequence 10, Appl
855	33	22.4	333	1	US-08-436-463-6	Sequence 6, Appl	928	33	22.4	533	1	US-08-458-181-10	Sequence 10, Appl
856	33	22.4	333	2	US-08-024-253-6	Sequence 6, Appl	929	33	22.4	533	4	PCT-US93-02172-10	Sequence 10, Appl
857	33	22.4	333	2	US-08-997-080-36	Sequence 36, Appl	930	33	22.4	553	2	US-08-793-229-36	Sequence 36, Appl
858	33	22.4	333	2	US-08-997-362-36	Sequence 36, Appl	931	33	22.4	553	3	US-09-285-957-36	Sequence 36, Appl
859	33	22.4	333	3	US-08-873-970-36	Sequence 36, Appl	932	33	22.4	560	4	PCT-US94-07644A-2	Sequence 2, Appl
860	33	22.4	334	2	US-08-997-080-45	Sequence 45, Appl	933	33	22.4	577	2	US-08-852-153-4	Sequence 4, Appl
861	33	22.4	334	2	US-08-997-362-45	Sequence 45, Appl	934	33	22.4	591	2	US-08-852-153-6	Sequence 6, Appl
862	33	22.4	334	2	US-08-873-970-45	Sequence 45, Appl	935	33	22.4	620	2	US-08-852-153-2	Sequence 2, Appl
863	33	22.4	338	2	US-08-447-430A-39	Sequence 39, Appl	936	33	22.4	627	2	US-08-466-589-6	Sequence 6, Appl
864	33	22.4	338	2	US-08-107-676-28	Sequence 28, Appl	937	33	22.4	627	2	US-08-700-636-6	Sequence 6, Appl
865	33	22.4	338	2	US-08-997-080-32	Sequence 32, Appl	938	33	22.4	627	3	US-08-467-574-6	Sequence 6, Appl
866	33	22.4	338	2	US-08-997-080-34	Sequence 34, Appl	939	33	22.4	652	2	US-08-751-305-2	Sequence 2, Appl
867	33	22.4	338	2	US-08-997-362-32	Sequence 32, Appl	940	33	22.4	659	1	US-08-258-639A-4	Sequence 4, Appl
868	33	22.4	338	2	US-08-997-362-34	Sequence 34, Appl	941	33	22.4	659	2	US-08-900-951-4	Sequence 4, Appl
869	33	22.4	338	3	US-08-873-970-32	Sequence 32, Appl	942	33	22.4	659	4	PCT-US95-07391A-4	Sequence 4, Appl
870	33	22.4	338	3	US-08-873-970-34	Sequence 34, Appl	943	33	22.4	664	1	US-08-421-661-6	Sequence 6, Appl
871	33	22.4	340	1	US-08-462-195-2	Sequence 2, Appl	944	33	22.4	664	2	US-08-852-153-8	Sequence 8, Appl
872	33	22.4	340	1	US-08-636-883-4	Sequence 4, Appl	945	33	22.4	683	5	5210183-3	Patent No. 5210183
873	33	22.4	340	2	US-08-636-883-2	Sequence 2, Appl	946	33	22.4	686	2	US-08-849-480A-4	Sequence 4, Appl
874	33	22.4	340	2	US-08-636-883-4	Sequence 4, Appl	947	33	22.4	730	1	US-08-121-713B-58	Sequence 58, Appl
875	33	22.4	340	2	US-08-107-676-3	Sequence 3, Appl	948	33	22.4	730	1	US-08-835-268-58	Sequence 58, Appl
876	33	22.4	340	2	US-08-107-676-31	Sequence 31, Appl	949	33	22.4	730	2	US-09-060-692-58	Sequence 58, Appl
877	33	22.4	340	2	US-08-997-080-37	Sequence 37, Appl	950	33	22.4	730	3	US-08-833-391-58	Sequence 58, Appl
878	33	22.4	340	2	US-08-997-362-37	Sequence 37, Appl	951	33	22.4	730	4	PCT-US94-10151A-58	Sequence 58, Appl
879	33	22.4	340	3	US-08-873-970-37	Sequence 37, Appl	952	33	22.4	777	2	US-08-874-678-3	Sequence 3, Appl
880	33	22.4	344	5	5210183-2	Patent No. 5210183	953	33	22.4	826	1	US-07-638-431-2	Sequence 2, Appl
881	33	22.4	348	1	US-08-459-489-10	Sequence 10, Appl	954	33	22.4	826	4	PCT-US92-00018-2	Sequence 2, Appl
882	33	22.4	348	1	US-08-458-686-10	Sequence 10, Appl	955	33	22.4	860	1	US-08-092-817-4	Sequence 4, Appl
883	33	22.4	348	1	US-07-843-350C-10	Sequence 10, Appl	956	33	22.4	863	2	US-08-666-271-2	Sequence 2, Appl
884	33	22.4	350	2	US-08-852-481-2	Sequence 2, Appl	957	33	22.4	879	1	US-08-306-546C-2	Sequence 2, Appl
885	33	22.4	353	2	US-08-447-430A-35	Sequence 35, Appl	958	33	22.4	879	2	US-08-530-524A-2	Sequence 2, Appl
886	33	22.4	353	2	US-08-447-430A-37	Sequence 37, Appl	959	33	22.4	885	1	US-08-372-892-4	Sequence 4, Appl
887	33	22.4	356	1	US-08-228-162-2	Sequence 2, Appl	960	33	22.4	894	1	US-08-372-892-2	Sequence 2, Appl
888	33	22.4	356	2	US-08-827-190-5	Sequence 5, Appl	961	33	22.4	894	1	US-08-445-640-34	Sequence 34, Appl
889	33	22.4	365	2	US-08-724-974A-3	Sequence 3, Appl	962	33	22.4	894	3	US-08-170-558-34	Sequence 34, Appl
890	33	22.4	384	2	US-08-827-190-6	Sequence 6, Appl	963	33	22.4	895	1	US-08-123-161A-8	Sequence 8, Appl
891	33	22.4	388	1	US-08-290-448A-80	Sequence 80, Appl	964	33	22.4	895	1	US-08-483-278-8	Sequence 8, Appl
892	33	22.4	388	1	US-08-290-448A-80	Sequence 80, Appl	965	33	22.4	897	1	US-07-960-389-2	Sequence 2, Appl
893	33	22.4	388	1	US-08-175-069A-80	Sequence 80, Appl	966	33	22.4	911	3	US-09-074-579-4	Sequence 4, Appl
894	33	22.4	389	1	US-08-324-977-4	Sequence 4, Appl	967	33	22.4	932	3	US-08-968-752B-6	Sequence 6, Appl
895	33	22.4	389	1	US-08-324-977-10	Sequence 10, Appl	968	33	22.4	967	3	US-08-188-281B-13	Sequence 13, Appl
896	33	22.4	389	2	US-08-384-616-4	Sequence 4, Appl	969	33	22.4	967	4	PCT-US94-07280-13	Sequence 13, Appl
897	33	22.4	389	2	US-08-384-616-10	Sequence 10, Appl	970	33	22.4	967	4	PCT-US95-01087-13	Sequence 13, Appl
898	33	22.4	389	2	US-08-904-686A-4	Sequence 4, Appl	971	33	22.4	1005	2	US-08-935-450-2	Sequence 2, Appl
899	33	22.4	389	2	US-08-904-686A-10	Sequence 10, Appl	972	33	22.4	1021	1	US-07-910-760-12	Sequence 12, Appl
900	33	22.4	396	4	PCT-US92-06965A-7	Sequence 7, Appl	973	33	22.4	1021	1	US-08-440-519-12	Sequence 12, Appl
901	33	22.4	422	1	US-08-396-218-2	Sequence 2, Appl	974	33	22.4	1074	2	US-08-470-058-2	Sequence 2, Appl
902	33	22.4	422	1	US-08-760-116-2	Sequence 2, Appl	975	33	22.4	1074	3	US-09-037-188-2	Sequence 2, Appl
903	33	22.4	425	4	PCT-US95-03866-8	Sequence 8, Appl	976	33	22.4	1141	1	US-08-131-365B-54	Sequence 54, Appl
904	33	22.4	442	2	US-08-821-355A-5	Sequence 5, Appl	977	33	22.4	1141	2	US-08-668-123-54	Sequence 54, Appl

Db 1980 PEILDVPS-TVQKTPFVTHPG 1999

RESULT 6

US-09-016-366A-12
; Sequence 12, Application US/09016366A
; Patent No. 5955431
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; APPLICANT: Huang, Chifu
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,366A
; FILING DATE: January 30, 1998
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/037,090
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7093
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2386 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-016-366A-12

Query Match 33.7%; Score 49.5; DB 2; Length 2386;
Best Local Similarity 42.9%; Pred. No. 81;
Matches 9; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 2 pdvqdcpctqlqenpfsspg 22
|::|||::|||
Db 2008 PEILDVPS-TVQKTPFVTHPG 2027

RESULT 7

US-08-551-356-2
; Sequence 2, Application US/08551356
; Patent No. 5830700
; GENERAL INFORMATION:
; APPLICANT: Irani, Meher
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA

; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551,356
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/349,762
; FILING DATE:
; APPLICATION NUMBER: US/07/998,271
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Holly, Julie A
; REGISTRATION NUMBER: 33-246
; REFERENCE/DOCKET NUMBER: 92-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2446 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-551-356-2

Query Match 33.7%; Score 49.5; DB 2; Length 2446;
Best Local Similarity 42.9%; Pred. No. 83;
Matches 9; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 2 pdvqdcpctqlqenpfsspg 22
|::|||::|||
Db 2099 PEILDVPS-TVQKTPFVTHPG 2118

RESULT 8

PCT-US93-12687-2
; Sequence 2, Application PC/TUS9312687
; GENERAL INFORMATION:
; APPLICANT: Irani, Meher H.
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/12687
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/998,271
; FILING DATE: 31-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 92-26PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322

SECRET

STREET: TWO MIL

STREET: Two Militia Drive

STREET: Two Militia Drive


```

:
:
: FILING DATE: 28-SEP-1993
:
: ATTORNEY/AGENT INFORMATION:
:
: NAME: NORMAN F. OBLON
:
: REGISTRATION NUMBER: 24,618
:
: REFERENCE/DOCKET NUMBER: 18-909-0 PCT
:
: TELECOMMUNICATION INFORMATION:
:
: TELEPHONE: 703-413-3000
:
: TELEFAX: 703-413-2220
:
: TELEX: 248855 OPAT UR
:
: INFORMATION FOR SEQ ID NO: 1:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 530 amino acids
:
: TYPE: amino acid
:
: TOPOLOGY: linear
:
: MOLECULE TYPE: peptide
:
:

```


US-08-513-841-1

Query Match 30.3%; Score 44.5; DB 1; Length 530;
Best Local Similarity 52.0%; Pred. No. 76;
Matches 13: Conservative 0; Mismatches 9; Indels 3; Gaps 1;

Qy 4 vqdcpectlqenpf---sqpgapi 25
158 VOSCOEMGLPYNPDFNGASOEGAGI 182

RESULT 14
 US-08-696-834-1
 ; Sequence 1, Application US/08696834
 ; Patent No. 5834263
 ; GENERAL INFORMATION:
 ; APPLICANT: Niwa, Mineo
 ; APPLICANT: Saito, Yoshimasa
 ; APPLICANT: Ishii, Yoshinori
 ; APPLICANT: Yoshida, Masaru
 ; APPLICANT: Hayashi, Hiromi
 ; TITLE OF INVENTION: Method for Producing 2-Keto-L-Gulonic Acid
 ; NUMBER OF SEQUENCES: 48
 ; CORRESPONDENCE ADDRESS:
 ;

ADDRESS: Oblon, Spivak, McClelland, Maier & Neustadt,
P.C.
ADDRESS: P.C.
STREET: 1755 Jefferson Davis Highway, Suite 400
City: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22202

```

:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Diskette - 3.50 Inch, 1.44 Mb storage
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE:
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/596,834
:
: FILING DATE: 24-SEP-1996
:
: CLASSIFICATION: 435
:
: PRIOR APPLICATION DATA:
:
: APPLICATION NUMBER: JP 28612/1994
:
: FILING DATE: 25-FEB-1994
:
: ATTORNEY/AGENT INFORMATION:
:

```

```

,
,
, NAME:
,
,   REGISTRATION NUMBER:
,
,   REFERENCE/DOCKET NUMBER:
,
, TELECOMMUNICATION INFORMATION:
,
,   TELEPHONE: (703) 413-3000
,
,   TELEFAX: (703) 413-2220
,
,   TELEX: 248855 OPAT UR
,
, INFORMATION FOR SEQ ID NO: 1:
,
,   SEQUENCE CHARACTERISTICS:
,
,     LENGTH: 530 amino acids
,
,     TYPE: amino acid
,
,     TOPOLOGY: linear
,
,     MOLECULE TYPE: peptide
,
,     ORIGINAL SOURCE:
,
,     ORGANISM: Gluconobacter oxydans
,
,     STRAIN: T-100
,
,     FEATURE:
,
,   NAME/KEY: mat peptide
,
,     LOCATION: 1..530
,
,     IDENTIFICATION METHOD: experimentally
,
, US-08-696-834-1

```

Query Match 30.3%; Score 44.5; DB 2; Length 530;
Best Local Similarity 52.0%; Pred. No. 76;
Matches 13; Conservative 0; Mismatches 9; Indels

Oy 4 vqdcpectlqenpf---sqpgapi 25
 ||| | | | | |
Db 158 VQSCQEMGLPYNPDFNGASQEGAGI 182

```

RESULT 15
US-08-942-673-1
; Sequence 1, Application US/08942673
; Patent No. 5861292
; GENERAL INFORMATION:
; APPLICANT: Niwa, Mineo
; APPLICANT: Saito, Yoshimasa
; APPLICANT: Ishii, Yoshinori
; APPLICANT: Yoshida, Masaru
; APPLICANT: Suzuki, Hiromi
; TITLE OF INVENTION: No. 5861292
; TITLE OF INVENTION: L-sorbose
; TITLE OF INVENTION: oxydants T-1
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Obion, Spivak, MCC
; STREET: 1755 Jefferson Davis

```

CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS Editor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/942,673
FILING DATE:

```

> CLASSIFICATION:
>
> PRIOR APPLICATION DATA:
> APPLICATION NUMBER: 08/513,841
> FILING DATE: 01-NOV-1995
> APPLICATION NUMBER: UK 9304700.9
> FILING DATE: 08-MAR-1993
>
> APPLICATION DATA: JP 241851/1993
> FILING DATE: 28-SEP-1993
> ATTORNEY/AGENT INFORMATION:
> NAME: NORMAN F. OBLON
> REGISTRATION NUMBER: 24,618
> REFERENCE/DOCKET NUMBER: 18-909-0 PCT
> TELECOMMUNICATION INFORMATION:
> TELEPHONE: 703-413-3000
> TELEFAX: 703-413-2220
> TELEX: 248855 OPAT UR
> INFORMATION FOR SEQ ID NO: 1:
> SEQUENCE CHARACTERISTICS:
> LENGTH: 530 amino acids
> TYPE: amino acid
> TOPOLOGY: linear
> MOLECULE TYPE: peptide
> ORIGINAL SOURCE:
> ORGANISM: Gluconobacter oxydans
> STRAIN: T-100
> FEATURE:
> NAME/KEY: mat peptide
> LOCATION: 1..530
> IDENTIFICATION METHOD: experimentally
>
> US-08-942-673-1

```


NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: ATG-50038
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1013 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-991-408-2

Query Match 30.3%; Score 44.5; DB 3; Length 1013;
Best Local Similarity 57.1%; Pred. No. 1.6e+02;
Matches 12; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

QY 7 cpec--tlqe-npfsggap 24
||| ||||| |||
DB 346 CPACGETLQESNGNLSPGFP 366

RESULT 19
US-08-775-009-34
Sequence 34, Application US/08775009
Patent No. 5935783
GENERAL INFORMATION:
APPLICANT: Gong, Weilong
APPLICANT: Emanuel, Beverly S.
APPLICANT: Budarf, Marcia L.
APPLICANT: Roe, Bruce
TITLE OF INVENTION: No. 5935783el Genes Mapping in the Digeorge and
TITLE OF INVENTION: Veliocardiofacial Syndrome Minimal Critical Region
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 5935783ris, LLP
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/775,009
FILING DATE: 27-DEC-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Trujillo, Doreen Yanko
REGISTRATION NUMBER: 35,719
REFERENCE/DOCKET NUMBER: CH-0681
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 871 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-775-009-34

Query Match 29.9%; Score 44; DB 2; Length 871;

Best Local Similarity 43.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 3 dvqdcpectlqenpf 18
| : ||| | : || :
DB 121 DSNNSPECFLRENAY 136
RESULT 20
US-08-342-930-2
Sequence 2, Application US/08342930
Patent No. 5821084
GENERAL INFORMATION:
APPLICANT: OLMSTED, ELIZABETH A.
APPLICANT: MAURO, LAURA J.
APPLICANT: DAVIS, ALAN R.
APPLICANT: DIXON, JACK E.
TITLE OF INVENTION: OSTEOBLAST-TESTICULAR PROTEIN TYROSINE
TITLE OF INVENTION: PHOSPHATASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/342,930
FILING DATE: 21-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KOSKI, ANTOINETTE F.
REGISTRATION NUMBER: 34,202
REFERENCE/DOCKET NUMBER: 20344-20975.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1711 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-342-930-2

Query Match 29.9%; Score 44; DB 2; Length 1711;
Best Local Similarity 42.9%; Pred. No. 3.3e+02;
Matches 9; Conservative 4; Mismatches 6; Indels 2; Gaps 1;

QY 4 vqdcpectlqenpfsggap 24
| : ||| | : |||
DB 1485 VEECPEDSMLASLF--PGGP 1503

RESULT 21
US-08-469-486-52
Sequence 52, Application US/08469486
Patent No. 5739281
GENERAL INFORMATION:
APPLICANT: Thøgersen, Hans Christian
APPLICANT: Holtet, Thor Las
APPLICANT: Etzerodt, Michael
TITLE OF INVENTION: Improved method for the refolding of
TITLE OF INVENTION: proteins
NUMBER OF SEQUENCES: 58

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Richardson
;; STREET: 225 Franklin Street
;; CITY: Boston
;; STATE: Massachusetts
;; COUNTRY: USA
;; ZIP: 02110-2804
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version
;; SOFTWARE: #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/469,486
;; FILING DATE:
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/192,060
;; FILING DATE: February 4, 1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Paul T. Clark
;; REGISTRATION NUMBER: 30,162
;; REFERENCE/DOCKET NUMBER: 06363/002001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617 542 5070
;; TELEFAX: 617 542 8906
;; TELEX: 200154
;; INFORMATION FOR SEQ ID NO: 52:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4544 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-469-486-52

Query Match 29.9%; Score 44; DB 1; Length 4544;
Best Local Similarity 53.3%; Pred. No. 1e+03;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 3 dvqdcpectlqenpf 17
|:||||| | | |
Db 3444 DERDCPEVTCAPNQF 3458

RESULT 22
US-08-469-658-52
; Sequence 52, Application US/08459658
; Patent No. 5917018
; GENERAL INFORMATION:
; APPLICANT: Th egersen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Etzerodt, Michael
; TITLE OF INVENTION: IMPROVED METHOD FOR THE REFOOLDING OF
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,658

;; FILING DATE: June 5, 1995
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/192,060
;; FILING DATE: February 4, 1994
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Paul T. Clark
;; REGISTRATION NUMBER: 30,162
;; REFERENCE/DOCKET NUMBER: 06363/002002
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617 542 5070
;; TELEFAX: 617 542 8906
;; TELEX: 200154
;; INFORMATION FOR SEQ ID NO: 52:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4544 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-469-658-52

Query Match 29.9%; Score 44; DB 2; Length 4544;
Best Local Similarity 53.3%; Pred. No. 1e+03;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 3 dvqdcpectlqenpf 17
|:||||| | | |
Db 3444 DERDCPEVTCAPNQF 3458

RESULT 23
US-08-709-979A-3
; Sequence 3, Application US/08709979A
; Patent No. 5912157
; GENERAL INFORMATION:
; APPLICANT: Claus von der Osten
; APPLICANT: Martin Sch lein
; TITLE OF INVENTION: No. 5912157el Alkaline Cellulases
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5912157o No. 5912157disk of No. 5912157th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,979A
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4160.404-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-709-979A-3

APPLICANT: SVEN, HASTRUP
TITLE OF INVENTION: An Enzyme Capable of Degrading Cellulose
TITLE OF INVENTION: or Hemicellulose

RESULT 27
US-08-483-432-27
; Sequence 27, Application US/08483432
; Patent No. 5763254
; GENERAL INFORMATION:
; APPLICANT: Woeldike, Helle F.
; APPLICANT: Hagen, Frederick
; APPLICANT: Hjort, Carsten M.
; APPLICANT: Sven, Hastrup
; TITLE OF INVENTION: An Enzyme Capable of Degrading Cellulose
; TITLE OF INVENTION: or Hemicellulose
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NO. 57632540 No. 5763254disk of No. 5763254th America, Inc.
; STREET: 405 Lexington Avenue, 62nd Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483.432
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/361,920
; FILING DATE:
; APPLICATION NUMBER: US 07/940,860
; FILING DATE: 28-OCT-1992

RESULT 28
 US-08-494-168-2
 ? Sequence 2, Application US/08494168
 ? Patent No. 5731192
 ? GENERAL INFORMATION:
 ? APPLICANT: Readers, Stephen T.
 ? APPLICANT: Zhou, Jing
 ? TITLE OF INVENTION: Collagen COL4A6; Gene, Protein and Method
 ? TITLE OF INVENTION: of Detecting Collagen Deficiency
 ? NUMBER OF SEQUENCES: 10
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESSEE: Foley & Lardner
 ? STREET: 3000 K Street, N.W., Suite 500
 ? CITY: Washington, D.C.
 ? COUNTRY: USA
 ? ZIP: 20007-5109
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: Floppy disk
 ? COMPUTER: IBM PC compatible
 ? OPERATING SYSTEM: PC-DOS/MS-DOS
 ? SOFTWARE: PatentIn Release #1.0, Version #1.25
 ? CURRENT APPLICATION DATA:
 ? APPLICATION NUMBER: US/08/494,168
 ? FILING DATE:
 ? CLASSIFICATION: 435
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: US 08/112,465
 ? FILING DATE: 27-AUG-1993
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: SAXE, Bernhard D.
 ? REGISTRATION NUMBER: 28,665
 ? REFERENCE/DOCKET NUMBER: 40397/104/BABR
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: (202)672-5300
 ? TELEFAX: (202)672-5399
 ? TELEX: 904136
 ? INFORMATION FOR SEQ ID NO: 2:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 1694 amino acids
 ? TYPE: amino acid
 ? TOPOLOGY: linear
 ? MOLECULE TYPE: protein
 ? US-08-494-168-2


```

Query Match      29.3%; Score 43; DB 1; Length 1694;
Best Local Similarity 36.8%; Pred. No. 4.6e+02;
Matches 7; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 4 vqdcptclqenffsqg 22
: | | | | |
Db 1562 ISRCVCEAPSKPFCAPG 1580

RESULT 29
US-08-813-884-42
; Sequence 42, Application US/08813884C
; Patent No. 6001410
; GENERAL INFORMATION:
; APPLICANT: Bolen, Paul L.
; APPLICANT: Kossiakoff, Nicholas
; APPLICANT: Hawn, Regina
; APPLICANT: Scharf Jr., Lewis G.
; TITLE OF INVENTION: A FRUIT LIQUOR BEVERAGE CONTAINING RECOMBINANT MONELLIN
; FILE REFERENCE: A fruit liquor beverage...SCM...enhan
; CURRENT APPLICATION NUMBER: US/08/813,884C
; CURRENT FILING DATE: 1997-03-06
; EARLIER APPLICATION NUMBER: 60/022,597
; EARLIER FILING DATE: 1996-07-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: As cited in
; OTHER INFORMATION: Somoza et al. 1995. Chemical Senses Journal
; OTHER INFORMATION: 20:61-68.
; FEATURE:
; OTHER INFORMATION: xaa = Gly, Ser, Glu or Thr
; PUBLICATION INFORMATION:
; TITLE: The Taste-active Regions of Monellin, a Potently Sweet
; TITLE: Protein.
; JOURNAL: Chemical Senses Journal
; VOLUME: 20
; PAGES: 61-68
; DATE: 1995
US-08-813-884-42

Query Match      28.6%; Score 42; DB 3; Length 80;
Best Local Similarity 35.3%; Pred. No. 20;
Matches 6; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 4 vqdcptclqenffsq 20
: | | | | |
Db 38 IRPCMKTIVENEXFRE 54

RESULT 30
US-08-991-300-2
; Sequence 2, Application US/08991300
; Patent No. 5973225
; GENERAL INFORMATION:
; APPLICANT: D'OVIDIO, RENATO
; APPLICANT: PORCEDDU, ENRICO
; APPLICANT: MERCHITELLI, CINZIA
; APPLICANT: CARDELLI, LUISA ERCOLI
; TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF A GENE
; TITLE OF INVENTION: ENCODING A LOW MOLECULAR WEIGHT GLUTENIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.

Query Match      28.6%; Score 42; DB 2; Length 369;
Best Local Similarity 40.9%; Pred. No. 1.1e+02;
Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 5 qdcptclqenffsqg 26
: | | | | |
Db 85 QQLPFPSQQQPPFSQQQPVL 106

RESULT 31
US-08-545-860D-55
; Sequence 55, Application US/08545860D
; Patent No. 6040140
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1 Region
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESS: No. 6040140ris
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,860D
; FILING DATE: 07-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04496

```


; FILING DATE: 22-APR-1994
; PCT-APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10930
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/327,392
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/320,559
; FILING DATE: 11-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/062,443
; FILING DATE: 14-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/971,094
; FILING DATE: 30-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,839
; FILING DATE: 27-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/805,093
; FILING DATE: 11-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca Esq., Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1262
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1093 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-545-860D-55

Query Match 28.6%; Score 42; DB 3; Length 1093;
Best Local Similarity 56.2%; Pred. NO. 3.9e+02;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 8 pectlqenpfssqgga 23
| | | | | | | |
Db 1055 PVLTAQTNPFSLSGA 1070

RESULT 32
PCT-US94-04496-55
; Sequence 55, Application PC/TUS9404496
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESSEE: Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04496
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca Esq., Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1242
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1093 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; PCT-US94-04496-55

Query Match 28.6%; Score 42; DB 4; Length 1093;
Best Local Similarity 56.2%; Pred. NO. 3.9e+02;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 8 pectlqenpfssqgga 23
| | | | | | | |
Db 1055 PVLTAQTNPFSLSGA 1070

RESULT 33
US-08-166-316-2
; Sequence 2, Application US/08166316
; Patent No. 5504192
; GENERAL INFORMATION:
; APPLICANT: Gill, Gordon N.
; APPLICANT: Wu, Rui-Yun
; TITLE OF INVENTION: ENDOCYTIC CODE BINDING PROTEIN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/166.316
; FILING DATE: 13-DEC-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr., Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-3204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-166-316-2

RESULT 35
US-08-720-484A-4
; Sequence 4, Application US/08720484A
; Patent No. 5990281
; GENERAL INFORMATION:
; APPLICANT: Desauvage, Frederic
; APPLICANT: Rosenthal, Arnon

CLASSIFICATION C14
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 960296.93839

TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1013 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-866-650-3

Query Match 27.6%; Score 40.5; DB 2; Length 1013;
Best Local Similarity 52.4%; Pred. No. 5.8e+02;
Matches 11; Conservative 1; Mismatches 6; Indels 3; Gaps 2;

QY 7 cpec-tlqenp-ffsqpgap 24
||| |||||
Db 346 CPACGETLQESSGNLSPPGFP 366

RESULT 37
US-09-021-287-3
Sequence 3, Application US/09021287
Patent No. 5981717
GENERAL INFORMATION:
APPLICANT: Greenspan, Daniel S
APPLICANT: Takahara, Kazuhiko
APPLICANT: Hoffman, Guy G
TITLE OF INVENTION: Mammalian Tolloid-Like Protein
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/021,287
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/866,650
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 960296.93839
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1013 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-021-287-3

Query Match 27.6%; Score 40.5; DB 2; Length 1013;
Best Local Similarity 52.4%; Pred. No. 5.8e+02;
Matches 11; Conservative 1; Mismatches 6; Indels 3; Gaps 2;

QY 7 cpec-tlqenp-ffsqpgap 24
||| |||||
Db 346 CPACGETLQESSGNLSPPGFP 366

RESULT 38
US-08-484-438-7
Sequence 7, Application US/08484438
Patent No. 5811098
Patent No. 5811098 5780031
GENERAL INFORMATION:
APPLICANT: Plowman, Gregory D.
APPLICANT: Culouscou, Jean-Michel
APPLICANT: Shoyab, Mohammed
APPLICANT: Siegall, Clay B.
APPLICANT: Hellstr m, Ingegerd
APPLICANT: Hellstr m, Karl E.
TITLE OF INVENTION: HERA HUMAN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,438
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,442
FILING DATE: 14-OCT-1994
APPLICATION NUMBER: US 08/150,704
FILING DATE: 10-NOV-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/981,165
FILING DATE: 24-NOV-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1210 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-438-7

Query Match 27.6%; Score 40.5; DB 2; Length 1210;
Best Local Similarity 37.0%; Pred. No. 7.1e+02;
Matches 10; Conservative 3; Mismatches 11; Indels 3; Gaps 1;

QY 3 dvqdcpectlqenpfisqpg---apil 26
||| | : |||||
Db 1009 DVDDADEVLPQGGFFSSPSTSRTPLL 1035

RESULT 39
US-08-475-035-4
Sequence 4, Application US/08475035
Patent No. 5985553


```
;
; GENERAL INFORMATION:
; APPLICANT: KING, C. R.
; APPLICANT: KRAUS, MATTHIAS H.
; APPLICANT: ARONSON, STUART A.
; TITLE OF INVENTION: HUMAN GENE RELATED TO BUT DISTINCT FROM
; TITLE OF INVENTION: EGF RECEPTOR GENE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, 127 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,035
; FILING DATE: 7 Jun 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perryman, David G.
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.656
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1210 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-475-035-4

Query Match 27.6%; Score 40.5; DB 2; Length 1210;
Best Local Similarity 37.0%; Pred. No. 7.1e+02;
Matches 10; Conservative 3; Mismatches 11; Indels 3; Gaps 1;

QY 3 dvqdcptclqenpfssqpg---april 26
||| | : ||| | :|
Db 1009 DVVDADEYLIPQOGFFSSPSTRTPL 1035

RESULT 40
US-08-026-138E-4
; Sequence 4, Application US/08026138E
; Patent No. 5502166
; GENERAL INFORMATION:
; APPLICANT: Masayoshi MISHINA
; TITLE OF INVENTION: NOVEL PROTEINS AND GENES CODING THE SAME
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nishiohata Residence 1-107
; STREET: 5214, Nishiohata-machi
; CITY: Niigata-shi
; STATE: Niigata-ken
; COUNTRY: JAPAN
; ZIP: 951
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS v.5
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/026,138E
; FILING DATE: 26-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 39563/1992
```

```
;
; FILING DATE: 26-FEB-1992
; APPLICATION NUMBER: JP 173155/1992
; FILING DATE: 30-JUN-1992
; APPLICATION NUMBER: JP 215017/1992
; FILING DATE: 12-AUG-1992
; APPLICATION NUMBER: JP 303878/1992
; FILING DATE: 13-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hamburg, C. Bruce
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-4551
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 986-2340
; TELEFAX: (212) 953-7733
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1323 amino acids
; TYPE: amino acid
; STRANDEDNESS: single strand
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORGANISM: mouse
; TISSUE TYPE: brain
; PUBLICATION INFORMATION:
; AUTHORS: Masayoshi MISHINA
; TITLE: NOVEL PROTEINS AND GENES CODING THE SAME
; RELEVANT RESIDUES IN SEQ ID NO: 4: FROM 1 to 1323
; US-08-026-138E-4

Query Match 27.6%; Score 40.5; DB 1; Length 1323;
Best Local Similarity 41.7%; Pred. No. 7.9e+02;
Matches 10; Conservative 2; Mismatches 11; Indels 1; Gaps 1;

QY 3 dvqdcptclqenpf-fsqpgapi 25
||| | : ||| | :|
Db 425 DTQHLTVALEERPFVIVEPADPI 448

RESULT 41
US-08-231-193A-58
; Sequence 58, Application US/08231193A
; Patent No. 5849895
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
; TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES THEREFOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,193A
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/052,459
; FILING DATE: 20-APR-1993
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
```



```

;
; GENERAL INFORMATION:
; APPLICANT: Benner, Steven A.
; TITLE OF INVENTION: Predicting Folded Structures of Proteins
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven A. Benner
; STREET: Hadlaubstrasse 151
; CITY: Zurich
; STATE: none
; COUNTRY: Switzerland
; ZIP: (note: this is an international post code) CH-8092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/857,224B
; FILING DATE: 03/25/92
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA: none
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (International) 41 1 632 2830
; TELEFAX: (International) 41 1 262 2437
; TELEX: none
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 264
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: rabbit
; FEATURE: Protein kinase; Table 8 Column 12
; PUBLICATION INFORMATION:
; AUTHORS:
; AUTHORS: Hanks, S. K.
; AUTHORS: Quinn, A. M.
; AUTHORS: Hunter, T.
; TITLE: The protein kinase family
; JOURNAL: Science
; VOLUME: 241
; PAGES: 42-52
; DATE: 1988
; US-07-857-224B-11

Query Match 27.2%; Score 40; DB 2; Length 264;
Best Local Similarity 42.1%; Pred. No. 1.5e+02;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Qy 6 dpectlqenpfqsgap 24
Db 43 DDVECTWVKRVLALPGKP 61

RESULT 45
US-07-857-224B-12
; Sequence 12, Application US/07857224B
; Patent No. 5958784
; GENERAL INFORMATION:
; APPLICANT: Benner, Steven A.
; TITLE OF INVENTION: Predicting Folded Structures of Proteins
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven A. Benner
; STREET: Hadlaubstrasse 151
; CITY: Zurich
; STATE: none
; COUNTRY: Switzerland
; ZIP: (note: this is an international post code) CH-8092
; COMPUTER READABLE FORM:

```

```

;
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/857,224B
; FILING DATE: 03/25/92
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA: none
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (International) 41 1 632 2830
; TELEFAX: (International) 41 1 262 2437
; TELEX: none
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 264
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: bovine
; FEATURE: Protein kinase; Table 8 Column 13
; PUBLICATION INFORMATION:
; AUTHORS:
; AUTHORS: Hanks, S. K.
; AUTHORS: Quinn, A. M.
; AUTHORS: Hunter, T.
; TITLE: The protein kinase family
; JOURNAL: Science
; VOLUME: 241
; PAGES: 42-52
; DATE: 1988
; US-07-857-224B-12

Query Match 27.2%; Score 40; DB 2; Length 264;
Best Local Similarity 42.1%; Pred. No. 1.5e+02;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Qy 6 dpectlqenpfqsgap 24
Db 43 DDVECTWVKRVLALPGKP 61

RESULT 46
US-08-741-437-1
; Sequence 1, Application US/08741437
; Patent No. 5843665
; GENERAL INFORMATION:
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: A NOVEL HUMAN PYROPHOSPHATASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/741,437
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

```


NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0148 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY:
CLONE: Consensus
US-08-741-437-1

Query Match 27.2%; Score 40; DB 2; Length 289;
Best Local Similarity 47.1%; Pred. No. 1.7e+02;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 7 cpecltqenpfqsgpa 23
| | | | | | | | | |
DB 242 CMNTLSESPFKCDPDA 258

RESULT 47
US-09-134-593-1
Sequence 1, Application US/09134593
Patent No. 5981232
GENERAL INFORMATION:
APPLICANT: Hawkins, Phillip R.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: A NOVEL HUMAN PYROPHOSPHATASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,593
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/741,437
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0148 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY:
CLONE: Consensus
US-09-134-593-1

Query Match 27.2%; Score 40; DB 2; Length 289;
Best Local Similarity 47.1%; Pred. No. 1.7e+02;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 7 cpecltqenpfqsgpa 23
| | | | | | | | | |
DB 242 CMNTLSESPFKCDPDA 258

RESULT 48
5266464-2
Patent No. 5266464
APPLICANT: HOUSEY, GERARD
TITLE OF INVENTION: METHOD OF SCREENING FOR PROTEIN INHIBITORS
AND ACTIVATORS
NUMBER OF SEQUENCES: 3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/392,073
FILING DATE: 10-AUG-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 154,206
FILING DATE: 10-FEB-1988
SEQ ID NO: 2:
LENGTH: 671
5266464-2

Query Match 27.2%; Score 40; DB 5; Length 671;
Best Local Similarity 42.1%; Pred. No. 4.3e+02;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 6 dpecltqenpfqsgpa 24
| | | | | | | | | |
DB 382 DDVECTMVEKRVLPKGP 400

RESULT 49
US-08-497-025-3
Sequence 3, Application US/08497025
Patent No. 5646251
GENERAL INFORMATION:
APPLICANT: Ruegg, Curtis L.
APPLICANT: Rivas, Alberto
APPLICANT: Laus, Reiner
APPLICANT: Engleman, Edgar G.
TITLE OF INVENTION: ALLOREACTION-ASSOCIATED ANTIGEN (ARAG)
TITLE OF INVENTION: A NOVEL MEMBER OF THE IMMUNOGLOBULIN GENE SUPERFAMILY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Hourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/497,025
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/149,212
FILING DATE: 05-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 5490A-219

Search completed: July 14, 2000, 09:32:06
Job time: 2180 sec

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1021 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-497-025-3

Query Match 27.2%; Score 40; DB 1; Length 1021;
Best Local Similarity 41.2%; Pred. No. 6.9e+02;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 9 ectlqenpfqsgapi 25
Db 23 ETVQKGPLFRAEGYV 39

RESULT 50
PCT-US95-10661A-6
Sequence 6, Application PC/TUS9510661A
GENERAL INFORMATION:
APPLICANT: Washington University, et al.
TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10661A
FILING DATE: 16-AUG-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,791
FILING DATE: 25-AUG-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Irecartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: FP-59941/RFT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1848 amino acids
TYPE: amino acid
TOPOLOGY: unknown
PCT-US95-10661A-6

Query Match 27.2%; Score 40; DB 4; Length 1848;
Best Local Similarity 42.1%; Pred. No. 1.3e+03;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 5 qdcptclqenpfqsgpa 23
Db 1300 QEVPQVASQESPKQEPAA 1318

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 14, 2000, 09:33:17 ; Search time 53.88 Seconds
(without alignments)
28.292 Million cell updates/sec

Title: ALPHA-CHAIN
Perfect score: 147
Sequence: 1 apvdqcdpctqlqenpfisqpgapil 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 168808 seqs, 58629743 residues

Total number of hits satisfying chosen parameters: 168807

Minimum DB seq length: 0
Maximum DB seq length: 1000000

Post-processing: Minimum Match 0%
Maximum Match 99%
Listing first 1000 summaries

Database :

PIR_63:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95	64.6	120	2 A39555	glycoprotein hormo
2	92	62.6	38	2 A60583	glycoprotein hormo
3	91	61.9	120	2 A45585	glycoprotein hormo
4	91	61.9	120	2 I51241	pituitary glycopro
5	90	61.2	96	2 S74086	foliitropin alpha
6	85	57.8	120	1 TTBOA	glycoprotein hormo
7	85	57.8	120	1 UTSHA	glycoprotein hormo
8	84	57.1	120	1 TTRTA	glycoprotein hormo
9	84	57.1	120	1 TTMSA	glycoprotein hormo
10	82	55.8	96	2 A05096	glycoprotein hormo
11	82	55.8	120	1 UTPGA	glycoprotein hormo
12	80	54.4	120	2 S53062	glycoprotein hormo
13	78	53.1	96	1 TTHOA	glycoprotein hormo
14	78	53.1	120	2 G00021	chorionic gonadotr
15	74	50.3	96	2 PN0138	glycoprotein hormo
16	73	49.7	117	2 A37198	glycoprotein hormo
17	73	49.7	118	1 UTCAA	glycoprotein hormo
18	73	49.7	118	2 SL6762	gonadotropin alpha
19	73	49.7	118	2 A06026	glycoprotein hormo
20	73	49.7	118	2 A40554	glycoprotein hormo
21	73	49.7	119	2 I51229	gonadotropin alpha
22	72	49.0	93	2 S07091	glycoprotein hormo
23	72	49.0	108	2 A06027	glycoprotein hormo
24	60	40.8	23	2 D60583	glycoprotein hormo
25	53	36.1	97	2 S20287	lutropin alpha cha
26	51	34.7	28	2 B60583	glycoprotein hormo
27	51	34.7	533	2 T10216	hypothetical prote
28	50.5	34.4	2265	1 FNBO	fibronectin - bovi
29	50	34.0	262	2 S56594	hypothetical 30.1k

30	49.5	33.7	2386	1 FNHU	fibronectin precu
31	49.5	33.7	2477	2 S14428	fibronectin precu
32	49	33.3	278	1 S25690	hupJ protein - Rho
33	48	32.7	117	2 I50992	gonadotropin alpha
34	48	32.7	1420	2 B57062	SRB9 protein - yea
35	47	32.0	55	2 D64392	rubredoxin - Metha
36	47	32.0	114	2 T26821	hypothetical prote
37	47	32.0	122	1 RNBV29	DNA-directed RNA p
38	47	32.0	1113	2 T20004	hypothetical prote
39	46.5	31.6	242	2 T29699	hypothetical prote
40	46.5	31.6	260	2 A43821	homeotic protein H
41	46.5	31.6	746	2 B70359	transcription regu
42	46	31.3	227	2 T16990	ribonuclease (EC 3
43	46	31.3	288	1 P3VXPS	3a protein - peanu
44	45.5	31.0	1034	2 S63536	aminomethyltransfe
45	45.5	31.0	1037	2 S63535	aminomethyltransfe
46	45.5	31.0	1037	2 S40216	P protein - Flaver
47	45	30.6	157	1 W5WLR1	E5 protein - rhesu
48	45	30.6	164	2 G71427	hypothetical prote
49	45	30.6	1021	2 G75403	DNA topoisomerase
50	45	30.6	1553	2 T09361	hypothetical prote
51	44.5	30.3	318	3 JE0284	Mm-1 cell derived
52	44.5	30.3	380	2 T28888	cuticle collagen d
53	44.5	30.3	397	2 A23707	aminomethyltransfe
54	44.5	30.3	472	2 T04855	hypothetical prote
55	44.5	30.3	572	2 T29880	hypothetical prote
56	44.5	30.3	819	2 A53714	protein kinase (EC
57	44	29.9	55	2 E70593	probable ruba prot
58	44	29.9	393	2 B72455	probable threonine
59	44	29.9	577	2 T18646	hypothetical prote
60	44	29.9	585	2 I37216	caliccin - human (f
61	44	29.9	588	2 I46003	caliccin - bovine
62	44	29.9	1569	2 G02757	clathrin heavy cha
63	44	29.9	1640	2 T09522	clathrin heavy cha
64	44	29.9	1711	1 A55148	protein-tyrosine-p
65	44	29.9	2180	2 T29764	hypothetical prote
66	44	29.9	4543	1 A53102	alpha-2-macroglobu
67	44	29.9	4544	1 S02392	alpha-2-macroglobu
68	44	29.9	4545	1 S25111	alpha-2-macroglobu
69	43.5	29.6	520	2 I51556	recombination acti
70	43.5	29.6	636	2 A57178	ubiquitin-activati
71	43.5	29.6	707	2 JC2218	procollagen C-endo
72	43	29.3	117	2 T38295	hypothetical prote
73	43	29.3	147	2 A48194	thyrotropin beta c
74	43	29.3	341	2 S57958	malate dehydrogena
75	43	29.3	511	2 T40334	hypothetical prote
76	43	29.3	572	2 T22547	hypothetical prote
77	43	29.3	750	2 T04980	hypothetical prote
78	43	29.3	788	2 S43567	RoH10.7 protein (
79	43	29.3	914	2 B75017	alanyl-tRNA synth
80	43	29.3	915	2 C71455	alanine--tRNA liga
81	43	29.3	1119	2 T16720	hypothetical prote
82	42.5	28.9	254	2 T25073	hypothetical prote
83	42.5	28.9	370	2 A41891	basal body P-ring
84	42.5	28.9	570	2 T04836	probable serine/th
85	42.5	28.9	904	2 T38757	homolog of Drosoph
86	42.5	28.9	926	2 T38756	homolog of Drosoph
87	42	28.6	100	2 T36231	probable dehydroge
88	42	28.6	114	2 B60627	glycoprotein hormo
89	42	28.6	114	2 I51230	gonadotropin alpha
90	42	28.6	178	2 T37360	hypothetical prote
91	42	28.6	307	2 S30432	hypothetical prote
92	42	28.6	332	2 A57140	motB protein homol
93	42	28.6	347	1 DEPUMW	malate dehydrogena
94	42	28.6	354	2 B75355	hypothetical prote
95	42	28.6	384	2 T30723	hypothetical prote
96	42	28.6	381	1 B69030	conserved hypothet
97	42	28.6	384	2 H64505	hypothetical prote
98	42	28.6	425	2 S39803	ribosomal protein
99	42	28.6	426	2 T04527	hypothetical prote
100	42	28.6	442	2 S20051	hypothetical prote
101	42	28.6	471	1 S78072	serine-type carbox
102	42	28.6	475	1 I64131	DNA repair protein

103	42	28.6	479	1	B65051	hypothetical prote	176	40	27.2	109	2	S70242	hypothetical prote
104	42	28.6	523	2	S55328	serine-type carbox	177	40	27.2	112	2	S21196	lutropin beta chai
105	42	28.6	601	2	T18787	hypothetical prote	178	40	27.2	129	2	S39445	DNA-directed RNA p
106	42	28.6	603	2	T08955	hypothetical prote	179	40	27.2	145	3	T39752	hypothetical prote
107	42	28.6	612	2	T32368	hypothetical prote	180	40	27.2	159	2	D65070	hypothetical prote
108	42	28.6	684	2	T25603	hypothetical prote	181	40	27.2	180	2	T09063	hypothetical prote
109	42	28.6	684	2	T31659	hypothetical prote	182	40	27.2	261	2	S49962	hypothetical prote
110	42	28.6	1033	2	T38533	AF17 protein - hum	183	40	27.2	266	2	S11900	hypothetical prote
111	42	28.6	1131	2	T15787	hypothetical prote	184	40	27.2	283	2	D72398	conserved hypothet
112	42	28.6	6831	2	T27934	hypothetical prote	185	40	27.2	297	3	T30900	2-hydroxy-6-oxo-6-
113	42	28.6	6839	2	S57242	twitchin - Caenorh	186	40	27.2	311	2	T33449	hypothetical prote
114	42	28.6	7160	2	T27935	hypothetical prote	187	40	27.2	323	2	T06506	glutenin low molec
115	41.5	28.2	53	2	G31436	ovomucoid, third d	188	40	27.2	342	2	T40090	hypothetical prote
116	41.5	28.2	56	2	A61493	ovomucoid (PSII-ty	189	40	27.2	365	2	T10213	hypothetical prote
117	41.5	28.2	241	2	T22216	hypothetical prote	190	40	27.2	393	2	A5863	integrase - Strept
118	41.5	28.2	328	2	A64878	hypothetical prote	191	40	27.2	407	2	C69217	peptide chain rele
119	41.5	28.2	387	2	T69971	spore coat protein	192	40	27.2	421	2	T00955	hypothetical prote
120	41.5	28.2	425	2	T18592	hypothetical prote	193	40	27.2	423	2	T28655	hypothetical prote
121	41.5	28.2	431	2	I51189	IFAPA-400 - chicke	194	40	27.2	478	2	C71523	probable shikimate
122	41.5	28.2	559	1	A35029	t-plasminogen acti	195	40	27.2	515	2	T25267	hypothetical prote
123	41.5	28.2	600	2	T18593	hypothetical prote	196	40	27.2	530	2	B36921	methyl coenzyme M
124	41.5	28.2	658	2	T33568	hypothetical prote	197	40	27.2	537	2	T25268	embryonic nuclear
125	41.5	28.2	1606	2	T34073	paranemin - chicke	198	40	27.2	539	2	A40581	embryonic nuclear
126	41.5	28.2	2236	1	QZFF	rudimentary protei	199	40	27.2	556	2	A32466	numb protein - fru
127	41.5	28.2	2960	2	A45259	desmoyokin - human	200	40	27.2	572	2	G02845	zyxin - human
128	41	27.9	33	2	D44798	phnA - Streptococ	201	40	27.2	574	1	S11032	5'-nucleotidase (E
129	41	27.9	50	2	F75282	ribosomal protein	202	40	27.2	574	1	JX0269	5'-nucleotidase (E
130	41	27.9	138	1	THUB	thyrotropin beta c	203	40	27.2	574	1	JC2001	5'-nucleotidase (E
131	41	27.9	231	2	T70968	hypothetical prote	204	40	27.2	576	1	A35036	hypothetical prote
132	41	27.9	253	2	T20415	hypothetical prote	205	40	27.2	580	2	T33742	hypothetical prote
133	41	27.9	283	2	T29837	hypothetical prote	206	40	27.2	623	2	S55652	transcription cont
134	41	27.9	294	2	T29839	hypothetical prote	207	40	27.2	623	2	T40056	nadh-cytochrome p
135	41	27.9	286	2	I52326	adenocorticotropi	208	40	27.2	671	1	KIHUC1	protein kinase C (
136	41	27.9	298	2	JC4046	adenocorticotropi	209	40	27.2	671	1	KIRTC1	protein kinase C (
137	41	27.9	298	2	JC1448	collagen col-34 -	210	40	27.2	671	1	KIRBC1	protein kinase C (
138	41	27.9	335	2	T31559	hypothetical prote	211	40	27.2	673	1	KIHUC2	protein kinase C (
139	41	27.9	348	1	R5HUL3	ribosomal protein	212	40	27.2	673	1	KIRTC2	protein kinase C (
140	41	27.9	370	2	T36352	hypothetical prote	213	40	27.2	673	1	KIRBC2	protein kinase C (
141	41	27.9	395	2	S40472	mitogen-activated	214	40	27.2	673	1	KIBOC2	protein kinase C (
142	41	27.9	455	2	A55050	enigma - human	215	40	27.2	673	2	S11213	protein kinase C (
143	41	27.9	474	2	T40627	probable Ribonucle	216	40	27.2	678	1	S29123	NADPH--ferrihemopr
144	41	27.9	546	2	JC4798	seizure-related me	217	40	27.2	685	2	T32571	hypothetical prote
145	41	27.9	582	2	B45878	hypothetical prote	218	40	27.2	957	2	A49847	nitrite reductase
146	41	27.9	740	2	A71141	hypothetical prote	219	40	27.2	1021	2	I39207	leukocyte surface
147	41	27.9	740	2	A75011	hypothetical prote	220	40	27.2	1127	2	T21635	hypothetical prote
148	41	27.9	1056	2	S55151	probable membrane	221	40	27.2	1145	1	GNLJEV	pol polyprotein -
149	41	27.9	1271	2	T24008	hypothetical prote	222	40	27.2	1146	1	GNLJEW	pol polyprotein (c
150	41	27.9	1421	2	T02501	hypothetical prote	223	40	27.2	1146	1	GNLJ22	pol polyprotein -
151	41	27.9	1502	2	T42216	multidrug resistan	224	40	27.2	1165	2	T21636	hypothetical prote
152	41	27.9	1879	2	T19481	hypothetical prote	225	40	27.2	1717	1	A45558	epidermal growth f
153	40.5	27.6	91	2	B75601	hypothetical prote	226	40	27.2	1849	2	C41859	iga-specific metal
154	40.5	27.6	200	2	JE0172	ribonuclease T2 (E	227	40	27.2	3898	2	S58295	polyprotein - hog
155	40.5	27.6	310	2	S04358	lacC protein - Sta	228	39.5	26.9	218	2	T36016	probable electron
156	40.5	27.6	373	2	T27414	hypothetical prote	229	39.5	26.9	218	2	T26826	hypothetical prote
157	40.5	27.6	390	2	T38191	actin-like protein	230	39.5	26.9	220	2	S70530	fibrial protein p
158	40.5	27.6	411	2	T03154	DNA polymerase pro	231	39.5	26.9	231	2	H72234	uridylyate kinase
159	40.5	27.6	431	2	G36809	hypothetical prote	232	39.5	26.9	249	2	S74373	heat shock protein
160	40.5	27.6	482	2	H72274	glutamyl tRNA-Gln	233	39.5	26.9	256	2	T03994	host response prot
161	40.5	27.6	538	2	A57624	retinoblastoma pro	234	39.5	26.9	263	2	A64807	endonuclease VIII
162	40.5	27.6	552	2	S42253	probable N3L prote	235	39.5	26.9	311	2	T21563	hypothetical prote
163	40.5	27.6	1037	2	T05309	probable glycine d	236	39.5	26.9	335	2	T31560	hypothetical prote
164	40.5	27.6	1077	2	S70120	ATP-dependent RNA	237	39.5	26.9	335	2	T31561	hypothetical prote
165	40.5	27.6	1210	1	GQHUE	epidermal growth f	238	39.5	26.9	336	2	E59845	thiamin biosynthes
166	40.5	27.6	1256	2	T03096	CDO protein - rat	239	39.5	26.9	360	2	T26270	hypothetical prote
167	40.5	27.6	1274	2	S55050	cardiac myosin-bin	240	39.5	26.9	376	2	S55275	fibromodulin precu
168	40.5	27.6	1323	2	T18557	N-methyl-D-asparta	241	39.5	26.9	449	2	A55362	procollagen I C-pr
169	40.5	27.6	1333	2	S27224	N-methyl-D-asparta	242	39.5	26.9	532	2	T15354	hypothetical prote
170	40.5	27.6	1356	1	C45219	N-methyl-D-asparta	243	39.5	26.9	815	2	B56708	extracellular sign
171	40.5	27.6	1385	2	S43230	156K protein - pla	244	39.5	26.9	937	2	B69068	leucine--trna liga
172	40.5	27.6	1579	2	S25329	carboxypeptidase y	245	39.5	26.9	1042	2	T37876	probable (xeroderm
173	40.5	27.6	1743	2	T26859	hypothetical prote	246	39.5	26.9	1246	2	S56752	helicase SK12W - h
174	40.5	27.6	2092	2	S30026	genome polyprotein	247	39.5	26.9	1873	2	T30944	surface protein pr
175	40.5	27.6	2149	2	S18676	genome polyprotein	248	39.5	26.9	1993	2	T30902	sodium channel SCA

249	39	26.5	80	2	T23206	hypothetical prote	322	39	26.5	2160	2	T20241	hypothetical prote
250	39	26.5	135	2	S31682	inhibin beta-A cha	323	39	26.5	3084	1	MMMSA	laminin alpha-1 ch
251	39	26.5	191	1	A64237	hypothetical prote	324	39	26.5	5175	2	T20992	hypothetical prote
252	39	26.5	191	2	JQ1399	hypothetical 21.5K	325	39	26.5	5198	2	T20993	hypothetical prote
253	39	26.5	212	2	C64843	probable transcrip	326	38.5	26.2	54	2	C31439	ovomucoid, third d
254	39	26.5	213	2	F70080	hypothetical prote	327	38.5	26.2	56	2	B31441	hypothetical prote
255	39	26.5	217	2	B41256	sdr protein - chic	328	38.5	26.2	110	2	T02485	hypothetical prote
256	39	26.5	242	2	S69282	polypyrrolidone tra	329	38.5	26.2	118	2	T37269	hypothetical prote
257	39	26.5	246	2	S74559	peptidyl-prolyl ci	330	38.5	26.2	136	2	T27455	hypothetical prote
258	39	26.5	250	2	T30124	hypothetical prote	331	38.5	26.2	153	2	C71502	probable (3R)-hydr
259	39	26.5	265	2	S70247	hypothetical prote	332	38.5	26.2	153	2	S42836	F40F12.7 protein -
260	39	26.5	276	2	S70995	cell surface adhes	333	38.5	26.2	201	2	B69546	conserved hypothet
261	39	26.5	297	1	NBRT	apolipoprotein H p	334	38.5	26.2	207	2	F70748	hypothetical prote
262	39	26.5	297	2	C43265	adenocorticotropi	335	38.5	26.2	340	2	I51547	probable RNA-bindi
263	39	26.5	303	2	T25778	hypothetical prote	336	38.5	26.2	347	2	I51546	probable RNA-bindi
264	39	26.5	311	2	C75493	GTP-binding protei	337	38.5	26.2	350	2	B37316	xylylene monooxygena
265	39	26.5	319	2	E72009	pp-loop ATPase hom	338	38.5	26.2	353	2	C64187	outer membrane pro
266	39	26.5	326	2	B69412	transcription init	339	38.5	26.2	398	2	S77529	hypothetical prote
267	39	26.5	338	1	A47183	hemoglobin precurs	340	38.5	26.2	425	2	A45064	interferon consens
268	39	26.5	345	1	NBBO	apolipoprotein H p	341	38.5	26.2	464	2	C75470	isocitrate lyase -
269	39	26.5	365	2	S75244	potassium channel	342	38.5	26.2	504	2	T35817	probable Glu-tRNA(
270	39	26.5	367	1	D70395	conserved hypothet	343	38.5	26.2	508	2	T31924	hypothetical prote
271	39	26.5	389	2	T20604	hypothetical prote	344	38.5	26.2	586	2	T20009	hypothetical prote
272	39	26.5	397	2	S28274	hypothetical prote	345	38.5	26.2	745	2	T23893	hypothetical prote
273	39	26.5	402	3	T39493	ptr3 or ubiquitin-	346	38.5	26.2	789	2	I52701	K-cadherin - rat
274	39	26.5	416	1	JN0006	nerve growth facto	347	38.5	26.2	790	2	I37016	cadherin-6 - human
275	39	26.5	417	2	E30341	alpha-l-adrenergic	348	38.5	26.2	911	2	I56552	synapse-associated
276	39	26.5	427	2	A64945	alpha-l-adrenergic	349	38.5	26.2	946	1	A29550	methylenetetrahydr
277	39	26.5	429	2	D75526	probable membrane	350	38.5	26.2	991	2	B71315	conserved hypothet
278	39	26.5	466	1	P2WL31	L2 protein - human	351	38.5	26.2	1064	1	S74540	protein-tyrosine k
279	39	26.5	475	2	S76734	hypothetical prote	352	38.5	26.2	1200	2	T00042	BAH-protocadherin p
280	39	26.5	515	2	A40491	alpha-l-adrenergic	353	38.5	26.2	1396	1	VCBE40	major capsid prote
281	39	26.5	515	2	JC1525	alpha-lB-adrenergi	354	38.5	26.2	1487	1	ECBE41	immediate-early pr
282	39	26.5	516	2	JC2332	alpha-lB-adrenergi	355	38.5	26.2	1487	1	ECBEF6	155K transcription
283	39	26.5	517	2	A45121	alpha-lB-adrenergi	356	38.5	26.2	1690	1	CGHULB	collagen alpha 4(I
284	39	26.5	534	2	T04663	hypothetical prote	357	38.5	26.2	1840	2	T29091	transitin - chicke
285	39	26.5	568	1	A55377	CSE-binding protei	358	38.5	26.2	2027	2	S60123	hypothetical prote
286	39	26.5	576	2	S48354	hypothetical prote	359	38.5	26.2	2769	1	URBO	thyroglobulin prec
287	39	26.5	587	1	B44276	coat protein VP1 -	360	38.5	26.2	4957	2	T03455	ALR protein - huma
288	39	26.5	600	2	T41165	hypothetical prote	361	38.5	26.2	5262	2	T03454	ALR protein - huma
289	39	26.5	643	2	G72320	penicillin-binding	362	38	25.9	60	2	D70593	probable rubB prot
290	39	26.5	655	2	F69121	conserved hypothet	363	38	25.9	103	2	S62052	hypothetical prote
291	39	26.5	658	2	E59228	ATP synthase, subu	364	38	25.9	106	2	E71618	Zn-ribon transcri
292	39	26.5	682	1	RNRZC1	DNA-directed RNA p	365	38	25.9	121	2	S18067	thyroid hormone re
293	39	26.5	683	1	RNZM51	DNA-directed RNA p	366	38	25.9	122	2	S40152	hypothetical prote
294	39	26.5	691	1	S76521	hypothetical prote	367	38	25.9	137	2	T15522	hypothetical prote
295	39	26.5	718	2	H75485	hypothetical prote	368	38	25.9	153	2	T40332	probable glutaredo
296	39	26.5	729	2	A69202	conserved hypothet	369	38	25.9	166	2	T40988	hypothetical prote
297	39	26.5	735	2	T13646	hypothetical prote	370	38	25.9	188	2	A75392	hypothetical prote
298	39	26.5	750	2	D65051	hypP protein - Esc	371	38	25.9	203	2	S76108	hypothetical prote
299	39	26.5	770	2	T00204	LDL receptor relat	372	38	25.9	221	2	T31620	hypothetical prote
300	39	26.5	770	2	T00203	LDL receptor-relat	373	38	25.9	258	2	JC6307	homeobox protein B
301	39	26.5	793	2	JC5539	Smoothed protein	374	38	25.9	267	2	T14824	alpha kafirin - so
302	39	26.5	822	2	S19947	fibroblast growth	375	38	25.9	267	2	T14825	alpha kafirin - so
303	39	26.5	822	2	B49151	fibroblast growth	376	38	25.9	276	2	S38629	MyoD protein - ral
304	39	26.5	848	2	A54740	interleukin-4-indu	377	38	25.9	280	2	T33519	hypothetical prote
305	39	26.5	853	1	CHRTD1	potassium channel	378	38	25.9	286	2	T03434	probable transport
306	39	26.5	857	2	S56529	potassium channel	379	38	25.9	298	2	T06980	glutenin low molec
307	39	26.5	882	2	S58881	mutS protein homol	380	38	25.9	300	2	T33232	hypothetical prote
308	39	26.5	986	2	T03760	alpha-dextrin endo	381	38	25.9	314	2	T09683	gibberellin 7-oxid
309	39	26.5	1021	2	T05901	hypothetical prote	382	38	25.9	319	2	T25116	hypothetical prote
310	39	26.5	1021	2	S64506	protein kinase RUB	383	38	25.9	325	2	T24689	hypothetical prote
311	39	26.5	1039	2	T22117	hypothetical prote	384	38	25.9	327	2	A46484	apoptosis-mediatin
312	39	26.5	1064	2	H64516	hypothetical prote	385	38	25.9	327	2	B72511	hypothetical prote
313	39	26.5	1131	2	J00382	phytochrome A - ma	386	38	25.9	339	2	G75529	peptide ABC transp
314	39	26.5	1139	2	B70934	hypothetical prote	387	38	25.9	341	2	T41286	malate dehydrogena
315	39	26.5	1156	2	T23308	hypothetical prote	388	38	25.9	353	2	T26164	hypothetical prote
316	39	26.5	1222	2	S40977	hypothetical prote	389	38	25.9	360	1	JH0692	protein-tyrosine-p
317	39	26.5	1400	2	T31555	hypothetical prote	390	38	25.9	364	2	A34867	arrestin - fruit f
318	39	26.5	1435	2	A37793	erythrocyte-bindin	391	38	25.9	367	2	B11847	hypothetical prote
319	39	26.5	1511	2	A53151	pleiotropic drug r	392	38	25.9	369	2	T78877	cAMP responsive el
320	39	26.5	1527	2	JE0336	canalicular multir	393	38	25.9	381	2	T15246	hypothetical prote
321	39	26.5	2018	2	T34274	hypothetical prote	394	38	25.9	384	2	JC4629	somatostatin recep

395	38	25.9	404	2	T16272	hypothetical prote	468	38	25.9	1788	2	T29043	hypothetical prote
396	38	25.9	409	2	T29517	hypothetical prote	469	38	25.9	1852	2	A37860	calcium channel pr
397	38	25.9	413	2	JC4862	activin beta-A cha	470	38	25.9	1895	2	T15881	hypothetical prote
398	38	25.9	413	2	S48756	transcription fact	471	38	25.9	2064	2	T13707	still life protein
399	38	25.9	413	2	S04655	transcription fact	472	38	25.9	2095	2	T29529	genome polyprotein
400	38	25.9	413	2	T15326	hypothetical prote	473	38	25.9	2437	2	S42612	transmembrane prot
401	38	25.9	422	2	A26328	cyclin A - Atlanti	474	38	25.9	2476	3	T34022	notch protein - pig
402	38	25.9	436	2	S43613	transposase (inser	475	38	25.9	2355	2	A40043	notch protein homo
403	38	25.9	437	2	T10010	probable serine/th	476	38	25.9	2606	2	T03159	large tegument pro
404	38	25.9	444	2	A55071	hydrogen peroxide-	477	38	25.9	2813	1	VWU	von Willebrand fac
405	38	25.9	450	2	T18781	hypothetical prote	478	38	25.9	2948	2	T22664	hypothetical prote
406	38	25.9	453	2	A35642	alpha-2B-adrenengi	479	38	25.9	3002	2	A47221	fibrillin 1 precu
407	38	25.9	465	2	T41830	LEF-4 orf90 - Bomb	480	38	25.9	3033	1	GNWJ8	genome polyprotein
408	38	25.9	473	2	T13433	pyrophosphate-depe	481	38	25.9	3033	2	JF1303	polyprotein - hepa
409	38	25.9	475	2	T18878	CAMP responsive el	482	38	25.9	3951	1	VF1HB1	F1 protein - avian
410	38	25.9	489	2	T35259	probable L-asparag	483	37.5	25.5	48	1	S00130	elastase inhibitor
411	38	25.9	496	2	S68452	apoptosis inhibito	484	37.5	25.5	52	2	B61589	ovomucoid (PSRI-ty
412	38	25.9	497	2	S69545	apoptosis inhibito	485	37.5	25.5	54	2	C31443	ovomucoid, third d
413	38	25.9	508	2	A45477	CAMP response elem	486	37.5	25.5	54	2	I31446	ovomucoid, third d
414	38	25.9	509	2	JH0501	zinc finger protei	487	37.5	25.5	54	2	F31439	ovomucoid, third d
415	38	25.9	542	2	S76358	GMP synthase (glut	488	37.5	25.5	54	2	C31447	ovomucoid, third d
416	38	25.9	586	1	JVEPB7	DNA maturase B - p	489	37.5	25.5	54	2	I61588	ovomucoid (PSRI-ty
417	38	25.9	588	2	T25248	hypothetical prote	490	37.5	25.5	54	2	A61589	ovomucoid (PSRI-ty
418	38	25.9	596	2	S33540	catechol oxidase (491	37.5	25.5	56	2	E31440	ovomucoid 2, third
419	38	25.9	598	2	A46225	Nur-related transc	492	37.5	25.5	162	2	T41902	hypothetical prote
420	38	25.9	599	2	T07097	probable catechol	493	37.5	25.5	162	2	T42001	hypothetical prote
421	38	25.9	600	2	T06292	hypothetical prote	494	37.5	25.5	180	2	S43791	PBDX protein - hum
422	38	25.9	602	2	S23991	conjugative transf	495	37.5	25.5	236	2	A53853	apob mRNA editing
423	38	25.9	606	2	JW0054	amiloride-sensitiv	496	37.5	25.5	238	1	A34378	endothelin 3 precu
424	38	25.9	615	1	VCMVJA	env polyprotein pr	497	37.5	25.5	285	2	I77964	SP-10 - western ba
425	38	25.9	645	2	T39614	kinase-binding pro	498	37.5	25.5	299	2	SS4087	probable membrane
426	38	25.9	650	2	F72540	hypothetical prote	499	37.5	25.5	303	2	S77235	hypothetical prote
427	38	25.9	666	2	I58169	senaphorin III - m	500	37.5	25.5	321	2	S55640	hypothetical prote
428	38	25.9	675	2	G71640	NADH dehydrogenase	501	37.5	25.5	357	2	S18235	omega sccalin prec
429	38	25.9	692	2	A70933	probable nr2z prot	502	37.5	25.5	358	2	G02670	IL8-related recept
430	38	25.9	727	2	S42834	F40f12.5 protein -	503	37.5	25.5	371	2	F70555	G protein-coupled
431	38	25.9	739	2	S32948	hyfF protein - Rho	504	37.5	25.5	375	2	JC5069	hypothetical prote
432	38	25.9	747	2	I51579	complement factor	505	37.5	25.5	421	2	T27311	interferon consens
433	38	25.9	748	2	I48744	senaphorin A - mou	506	37.5	25.5	424	2	A35861	hypothetical protei
434	38	25.9	761	2	H65083	glycolate oxidase	507	37.5	25.5	449	2	D69037	PET112-like protei
435	38	25.9	766	2	G71160	probable histidine	508	37.5	25.5	476	2	S64953	hypothetical prote
436	38	25.9	771	2	D49423	senaphorin III pre	509	37.5	25.5	478	2	JK0201	alpha-amylose (EC
437	38	25.9	772	2	A90969	collapsin - chicke	510	37.5	25.5	484	2	B70586	probable amIA2 pro
438	38	25.9	772	2	I48747	senaphorin D - mou	511	37.5	25.5	498	2	A48305	alpha-amylose (EC
439	38	25.9	779	2	G70926	probable fdhf prot	512	37.5	25.5	499	1	ALAS1	alpha-amylose (EC
440	38	25.9	796	2	S78444	lethal(3)malignant	513	37.5	25.5	499	1	ALAS3	alpha-amylose (EC
441	38	25.9	805	2	A69682	primosomal replica	514	37.5	25.5	499	2	B48305	alpha-amylose (EC
442	38	25.5	812	2	T16621	hypothetical prote	515	37.5	25.5	499	2	JS0663	alpha-amylose (EC
443	38	25.9	819	1	B72128	endopeptidase La (516	37.5	25.5	499	2	JT0466	alpha-amylose (EC
444	38	25.9	828	2	S56250	probable membrane	517	37.5	25.5	499	2	JN0588	alpha-amylose (EC
445	38	25.9	862	1	QRMSLD	LDL receptor precu	518	37.5	25.5	522	2	S71821	probable interleuk
446	38	25.9	865	1	ISECTP	DNA topoisomerase	519	37.5	25.5	534	2	T33772	hypothetical prote
447	38	25.9	867	2	H70141	alanine--tRNA liga	520	37.5	25.5	568	1	UKHUT	t-plasminogen acti
448	38	25.9	925	2	T01384	hypothetical prote	521	37.5	25.5	568	1	I61106	involucrin - rat
449	38	25.9	929	2	T17392	vrl1 protein - Dic	522	37.5	25.5	590	2	S16411	terminase ATPase c
450	38	25.9	1017	2	T37201	hypothetical prote	523	37.5	25.5	640	1	VCNVRV	env polyprotein pr
451	38	25.9	1067	2	T06312	hypothetical prote	524	37.5	25.5	643	1	S15616	E1 protein - human
452	38	25.9	1072	2	T30846	serine-repeat anti	525	37.5	25.5	737	2	S47857	basic protein, cyt
453	38	25.9	1087	2	T30844	serine-repeat anti	526	37.5	25.5	755	2	S23441	hypothetical prote
454	38	25.9	1089	2	T30843	serine-repeat anti	527	37.5	25.5	789	2	S28259	androgen-regulated
455	38	25.9	1171	1	QOKBFP	pyruvate (flavodox	528	37.5	25.5	992	2	S04781	exonuclease ABC c
456	38	25.9	1308	2	E71622	probable membrane	529	37.5	25.5	1069	2	T00040	BH-protocadherin P
457	38	25.9	1311	2	T33757	hypothetical prote	530	37.5	25.5	1069	2	T00043	BH-protocadherin-a
458	38	25.9	1323	2	T11661	phosphoribosylform	531	37.5	25.5	1072	2	T00041	BH-protocadherin P
459	38	25.9	1346	3	T17412	polyketide synthas	532	37.5	25.5	1191	2	S27329	DNA topoisomerase
460	38	25.9	1354	2	T13363	phosphoribosylform	533	37.5	25.5	1192	1	ISXFA5	DNA topoisomerase
461	38	25.9	1361	2	T30884	neural specific DN	534	37.5	25.5	1197	1	A48350	DNA-binding protei
462	38	25.9	1369	2	T03104	tegument protein h	535	37.5	25.5	1210	2	A53183	epidermal growth f
463	38	25.9	1459	2	T24088	hypothetical prote	536	37.5	25.5	1218	2	T14265	epiglin-245 - mouse
464	38	25.9	1675	1	LRRTH	clathrin heavy cha	537	37.5	25.5	1231	2	T30841	serine-repeat anti
465	38	25.9	1681	2	S42369	Clathrin heavy cha	538	37.5	25.5	1298	2	A64157	hypothetical prote
466	38	25.9	1691	1	S22917	collagen alpha 5(I	539	37.5	25.5	1502	2	T31671	Rab3 GDP/GTP excha
467	38	25.9	1766	2	A42155	trophozoite cystei	540	37.5	25.5	25	2	D37264	E2 glycoprotein -

833	36	24.5	411	2	T29475	hypothetical prote	906	36	24.5	782	2	T15060	hypothetical prote
834	36	24.5	414	2	PSQ323	von Willebrand fac	907	36	24.5	785	1	QQV28	D5 protein - vacci
835	36	24.5	418	2	S27967	pigment epithelial	908	36	24.5	785	2	E72162	F5R protein - vari
836	36	24.5	418	2	A47281	pigment epithelium	909	36	24.5	786	2	A48456	ocyst wall protei
837	36	24.5	430	2	S77393	cell division prot	910	36	24.5	796	2	T36308	probable integral
838	36	24.5	431	1	E70699	probable pknA prot	911	36	24.5	802	2	JH0595	potassium channel
839	36	24.5	432	2	T32044	hypothetical prote	912	36	24.5	819	1	C71527	endopeptidase La (
840	36	24.5	443	2	T33690	hypothetical prote	913	36	24.5	836	2	D64200	DNA topoisomerase
841	36	24.5	459	1	GLAG	beta-glucosidase (914	36	24.5	839	2	A56337	glycoprotein phosp
842	36	24.5	461	2	JC2477	calcitonin recepto	915	36	24.5	840	2	T36175	probable large ATP
843	36	24.5	462	2	D69337	aspartate kinase (916	36	24.5	857	2	A42861	protein kinase TTK
844	36	24.5	464	2	I60194	calcitonin-like re	917	36	24.5	878	2	A69584	alanine--trna liga
845	36	24.5	469	1	P2WL35	L2 protein - human	918	36	24.5	919	2	T07391	probable alpha-glu
846	36	24.5	469	2	S36525	L2 protein - human	919	36	24.5	943	2	A42681	centromere protein
847	36	24.5	470	2	S57264	L-aminoadipate-sem	920	36	24.5	963	2	T19140	hypothetical prote
848	36	24.5	471	2	JH0533	tryptophan--trna l	921	36	24.5	976	2	A36355	protein-tyrosine k
849	36	24.5	471	2	A41706	tryptophan--trna l	922	36	24.5	996	2	JE0237	apolipoprotein E r
850	36	24.5	475	1	YWRBPR	tryptophan--trna l	923	36	24.5	997	2	A60776	230k bullous pemph
851	36	24.5	479	1	IXBE1F	alpha trans-induci	924	36	24.5	1014	2	T30840	serine-repeat anti
852	36	24.5	481	2	S50053	tryptophan--trna l	925	36	24.5	1015	2	T15830	hypothetical prote
853	36	24.5	488	2	T31534	hypothetical prote	926	36	24.5	1017	2	B70985	probable polyketid
854	36	24.5	490	1	IXBE17	alpha trans-induci	927	36	24.5	1020	2	JN0124	glycine dehydrogen
855	36	24.5	495	2	T33463	hypothetical prote	928	36	24.5	1020	2	B39521	glycine dehydrogen
856	36	24.5	496	2	G72289	glucose-6-phosphat	929	36	24.5	1023	2	JC4013	major acidic nucle
857	36	24.5	498	2	JC5983	protein kinase C-i	930	36	24.5	1035	2	I58409	integrin alpha-9 c
858	36	24.5	500	2	T06332	ferric leghemoglob	931	36	24.5	1059	2	T12195	sucrose-phosphate
859	36	24.5	505	2	C64502	hypothetical prote	932	36	24.5	1061	1	S27311	ribonuclease E (EC
860	36	24.5	512	2	T33463	probable serine ca	933	36	24.5	1100	2	H71616	SERA antigen/papal
861	36	24.5	514	2	T21286	hypothetical prote	934	36	24.5	1163	2	S07137	DNA-directed RNA p
862	36	24.5	515	1	S65578	lanosterol 14alpha	935	36	24.5	1172	2	T00065	hypothetical prote
863	36	24.5	518	2	T25507	hypothetical prote	936	36	24.5	1186	3	T33754	O/E-1-associated z
864	36	24.5	536	2	T37840	probable vacuolar	937	36	24.5	1203	1	A47501	nitric-oxide synth
865	36	24.5	542	2	T29707	hypothetical prote	938	36	24.5	1207	1	EGHU	epidermal growth f
866	36	24.5	542	2	A44358	zyxin - chicken	939	36	24.5	1221	2	A49457	fibulin-2 precurs
867	36	24.5	555	2	I38656	lymphoid-restrict	940	36	24.5	1231	2	S20638	desmoyokin - mouse
868	36	24.5	561	2	T27318	hypothetical prote	941	36	24.5	1252	2	S36016	ocyst wall protei
869	36	24.5	566	2	S59382	hypothetical prote	942	36	24.5	1281	2	T15762	hypothetical prote
870	36	24.5	570	1	A38954	hydroxylamine oxid	943	36	24.5	1302	2	T00038	hypothetical prote
871	36	24.5	574	2	S28275	hypothetical prote	944	36	24.5	1435	2	S54697	DNA polymerase III
872	36	24.5	577	2	T01052	hypothetical prote	945	36	24.5	1451	2	D64203	DNA polymerase III
873	36	24.5	579	1	D70740	hypothetical prote	946	36	24.5	1475	2	T29809	hypothetical prote
874	36	24.5	582	2	T37473	hypothetical zinc	947	36	24.5	1480	2	S48440	poly(A)-specific r
875	36	24.5	587	2	JE0268	myosin - mouse	948	36	24.5	1482	2	T34010	hypothetical prote
876	36	24.5	597	2	A46050	thyroid/steroid re	949	36	24.5	1534	2	S59604	DNA (cytosine 5)-
877	36	24.5	597	2	T74612	translation elonga	950	36	24.5	1558	2	T34394	hypothetical prote
878	36	24.5	598	2	I84692	nuclear orphan rec	951	36	24.5	1607	1	MMHUB2	laminin gamma-1 ch
879	36	24.5	602	1	TVTRRR	protein kinase (EC	952	36	24.5	1609	1	MMHUB2	laminin gamma-1 ch
880	36	24.5	603	2	S55869	histidine transpor	953	36	24.5	1613	2	A43081	vitellogenin 2 pre
881	36	24.5	605	1	S07314	UDPGlucose--starch	954	36	24.5	1613	2	T15793	vitellogenin - Gae
882	36	24.5	608	2	T14731	UDPGlucose--starch	955	36	24.5	1715	2	JE0128	Bombyx mori recept
883	36	24.5	615	2	T36183	probable ABC-type	956	36	24.5	1720	2	T07258	cell division prot
884	36	24.5	617	2	T10102	pyrophosphate--fru	957	36	24.5	1783	2	T42386	unconventional myo
885	36	24.5	622	2	I48733	protein kinase rck	958	36	24.5	1799	1	JC4920	ZK688.5 protein -
886	36	24.5	626	2	T27319	hypothetical prote	959	36	24.5	1872	2	UC4976	plexin 3 precursor
887	36	24.5	629	2	E72297	methionine--trna l	960	36	24.5	1998	2	T13009	hypothetical prote
888	36	24.5	646	2	T27899	hypothetical prote	961	36	24.5	2140	2	T18543	probable cell-adhe
889	36	24.5	648	1	TVHUF6	protein kinase raf	962	36	24.5	2163	2	T15276	hypothetical prote
890	36	24.5	648	1	TVTRRF	protein kinase raf	963	36	24.5	2167	2	T34395	hypothetical prote
891	36	24.5	648	2	T31658	probable primosoma	964	36	24.5	2185	1	S60200	acetyl-CoA carboxy
892	36	24.5	649	2	T01882	hypothetical prote	965	36	24.5	2215	2	T16871	hypothetical prote
893	36	24.5	658	2	T35187	anaerobic ribonuci	966	36	24.5	2318	2	S45306	notch3 protein -
894	36	24.5	665	1	VCVWEM	env polyprotein -	967	36	24.5	2321	2	S78549	notch3 protein - h
895	36	24.5	665	2	T04290	hypothetical prote	968	36	24.5	2447	2	T16870	hypothetical prote
896	36	24.5	667	2	S56829	probable membrane	969	36	24.5	2524	2	A35844	Xotch protein - Af
897	36	24.5	667	2	A48579	trophozoite surfac	970	36	24.5	2616	2	A57096	nudel protein prec
898	36	24.5	681	2	C75395	methionyl-RNA syn	971	36	24.5	3418	1	G02334	breast cancer tumo
899	36	24.5	727	1	A45342	nonstructural prot	972	36	24.5	3623	2	T09456	intrinsic factor-B
900	36	24.5	747	2	T16274	hypothetical prote	973	36	24.5	3866	2	B48205	All-1 protein -GRE
901	36	24.5	758	2	T25751	hypothetical prote	974	36	24.5	3869	2	A48205	All-1 protein +GRE
902	36	24.5	770	2	S04847	leukocyte adhesion	975	36	24.5	4302	2	A38971	polycystic kidney
903	36	24.5	771	2	A45839	leukocyte adhesion	976	36	24.5	6486	3	T31076	tyrocytic synthe
904	36	24.5	780	2	A34102	von Willebrand fac	977	36	24.5	6558	2	T13931	projectin - fruit
905	36	24.5	780	2	A48143	HF-1 regulatory el	978	36	24.5	7962	2	I38346	elastic titin - hu

orf 5' of meg1 - m	38	2	B49012
ovomucoid (PSTI-ty	51	2	C61587
ovomucoid (PSTI-ty	54	2	A61494
ovomucoid (PSTI-ty	56	2	B61494
ovomucoid, third d	56	2	E31438
ovomucoid 1, third	56	2	D31440
ovomucoid, third d	56	2	G31440
hypothetical prote	111	2	D72646
Ig gamma chain - A	112	2	S33599
hypothetical prote	136	2	S77562
NADH dehydrogenase	172	2	S16208
thiol:disulfide in	185	2	F75549
interferon alpha-I	189	1	IYBO11
interferon alpha-I	189	1	IVBO1D
hypothetical prote	202	2	T05763
ovomucoid precurs	210	1	TIGDM
DNA-binding protei	225	2	S39852
hypothetical prote	237	2	T25877
homeotic protein H	250	2	A56572
hypothetical prote	263	2	T17302
hypothetical prote	285	2	T19689

ALIGNMENTS

RESULT 1
A39555
glycoprotein hormones alpha chain precursor - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C:date: 08-Nov-1991 #sequence_revision 08-Nov-1991 #text_change 13-Jan-1995
C:Accession: A39555
R:Giles, T. G. & Durning, M. & Fisher, J. M.
DNA: Cell, Biochem 367-380, 1991
Abstract: Molecular cloning of the rhesus glycoprotein hormone alpha-subunit gene.
A:Reference-number: A39555; MUID:91321740
A:Accession: A39555
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-120 <GOL>
C:Superfamily: glycoprotein hormones alpha chain
C:Keywords: glycoprotein

Matches 16; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 vqdcpectlqenffsqgap1 25
|| |||| | |||:||||:
Db 2 VQGCPECLGENFFSKPGAPV 23

RESULT 3
A45585
glycoprotein hormones alpha chain precursor - turkey
N;Alternate names: gonadotropin alpha chain; thyrotropin alpha chain
C;Species: Meleagris gallopavo (Common turkey)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 16-Jul-1999
C;Accession: A45585; A36692; C60583
R;Foster, D.N.; Foster, L.K.
Poult. Sci. 70, 2516-2523, 1991
A;Title: Cloning and sequence analysis of the common alpha-subunit complementary deox
A;Reference number: A45585; MUID:92150364
A;Accession: A45585
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-120 <FOSI>
A;Experimental source: pituitary
A;Note: Sequence extracted from NCBI backbone (NCBIN:82029, NCBIIP:82032)
Foster, D.N.; Kim, S.U.; Enyeart, J.J.; Foster, L.K.
Biochem Biophys Res Commun: 173, 967-973, 1990
A;Title: Nucleotide sequence of the complementary DNA for turkey growth hormone.
A;Reference number: A36692; MUID:91097605
A;Accession: A36692
A;Molecule type: mRNA
A;Residues: 1-120 <FOS2>
A;Cross-references: GB:M33698; NID:g213895; PIDN:AAA49629.1; PID:g213896
A;Note: authors inadvertently published glycoprotein hormones alpha chain sequence in
R;Bergert, E.R.; Madden, B.; McCormick, D.J.; Papkoff, H.; Ryan, R.J.
Endocrinology 127, 2985-2989, 1990
A;Title: The antigenic structure of the human glycoprotein hormone alpha-subunit: Cro
A;Reference number: A60583; MUID:91065269
A;Accession: C60583
A;Molecule type: protein
A;Residues: 34-61 <BER>
C;Superfamily: glycoprotein hormones alpha chain
C;Keywords: glycoprotein; hormone
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-120/Product: glycoprotein hormones alpha chain, minor form #status experimental
F;34-120/Product: glycoprotein hormones alpha chain, major form #status experimental

Query Match 61.9%; Score 91; DB 2; Length 120;
Best Local Similarity 64.3%; Pred. No. 7.4e-06;
Matches 18; Conservative 2; Mismatches 4; Indels 4; Gaps 1;
QY 2 pd----vqdcpectlqenffsqgap1 25
|| : ||||| ||| |||:||||
Db 26 PDGEFLMGCCPECLGENFFSKPGAPI 53

RESULT 4
I51241
Pituitary glycoprotein hormone alpha-subunit - quail
C;Species: Coturnix coturnix (quail)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C;Accession: I51241
R;Ando, H.; Ishii, S.
Gen. Comp. Endocrinol. 93, 357-368, 1994
A;Title: Molecular cloning of complementary deoxyribonucleic acids for the pituitary
urnix coturnix japonica).
A;Reference number: I51241; MUID:94252550
A;Accession: I51241
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-120 <AND>
A;Cross-references: GB:S70833; NID:g546919; PIDN:AAB30866.1; PID:g546920
C;Superfamily: glycoprotein.hormones alpha chain

541	37	25.2	52	2	G71256	probable rubredoxin	614	37	25.2	523	2	T08854	ferric leghemoglobin
542	37	25.2	53	1	JU0127	rubredoxin - "Buty	615	37	25.2	528	2	T24698	hypothetical prote
543	37	25.2	54	2	G36769	17R protein - huma	616	37	25.2	541	2	T19224	hypothetical prote
544	37	25.2	79	2	T39177	hypothetical prote	617	37	25.2	546	2	T19222	hypothetical prote
545	37	25.2	82	2	S02660	9K protein (clone	618	37	25.2	548	2	S59133	ET52 repressor fac
546	37	25.2	84	2	S74883	hypothetical prote	619	37	25.2	571	2	T2427	probable translati
547	37	25.2	119	2	A48230	T-cell receptor be	620	37	25.2	584	2	S75986	hypothetical prote
548	37	25.2	124	2	T15965	hypothetical prote	621	37	25.2	590	2	A40437	glutamic acid-rich
549	37	25.2	138	1	TTBOB	thyrotropin beta c	622	37	25.2	593	2	A71361	probable thiamin A
550	37	25.2	169	2	T26882	hypothetical prote	623	37	25.2	596	2	A45664	variant-specific s
551	37	25.2	180	2	JC4955	kappa-casein precu	624	37	25.2	605	2	T19818	hypothetical prote
552	37	25.2	207	1	XUHQMC	methylated-DNA--pr	625	37	25.2	606	2	T31557	hypothetical prote
553	37	25.2	212	2	E64123	stringent starvati	626	37	25.2	609	2	S69334	chloride peroxidase
554	37	25.2	213	2	T36895	probable 3-methyla	627	37	25.2	609	2	T32302	hypothetical prote
555	37	25.2	228	2	S53497	ribonuclease (EC 3	628	37	25.2	635	2	T41367	hypothetical prote
556	37	25.2	250	2	T22786	hypothetical prote	629	37	25.2	691	1	TVVPAS	large T antigen -
557	37	25.2	252	2	A34877	C4b-binding protei	630	37	25.2	695	1	TVVPTB	large T antigen -
558	37	25.2	253	2	E64025	hypothetical prote	631	37	25.2	716	2	G01627	androgen receptor
559	37	25.2	269	2	G70710	enoyl-lacyl-carrie	632	37	25.2	749	2	G01856	semaphorin V - hum
560	37	25.2	272	1	S39680	hydroxyethylthiaz	633	37	25.2	783	2	I50116	N-cadherin precurs
561	37	25.2	287	2	F65165	33.2 kD protein in	634	37	25.2	790	2	I51638	F-cadherin - Afric
562	37	25.2	289	2	A45153	inorganic pyrophos	635	37	25.2	803	2	T40036	hypothetical prote
563	37	25.2	296	1	G64453	conserved hypothet	636	37	25.2	863	2	S06017	neuraxin - rat
564	37	25.2	297	2	S42767	adrenocorticotropi	637	37	25.2	891	1	G75014	valyl-tRNA synthet
565	37	25.2	297	2	I45849	gene ACTH receptor	638	37	25.2	894	1	SVBYMX	leucine--tRNA liga
566	37	25.2	305	2	I38144	phosphotyrosine pi	639	37	25.2	940	2	A40985	proctectin - fruit
567	37	25.2	319	2	D75349	conserved hypothet	640	37	25.2	946	2	G71617	SERA antigen/papai
568	37	25.2	320	1	A39479	homeotic protein m	641	37	25.2	1034	2	JC5598	mucin - rat
569	37	25.2	321	1	S28390	homeotic protein m	642	37	25.2	1061	1	GNLJG4	pol polyprotein -
570	37	25.2	324	2	JC2395	Fas antigen precu	643	37	25.2	1072	2	T37742	serine threonine-p
571	37	25.2	325	2	B72098	dipeptidase - Chla	644	37	25.2	1148	2	T09073	splicing factor Si
572	37	25.2	327	2	JS0402	gamma-gliadin prec	645	37	25.2	1186	2	T19050	hypothetical prote
573	37	25.2	329	1	G2GP	Ig gamma-2 chain c	646	37	25.2	1217	2	T18209	mxg protein - Bac
574	37	25.2	333	2	S75980	hypothetical prote	647	37	25.2	1239	1	VHWVEE	structural polypro
575	37	25.2	335	2	JT0569	chondromodulin-I p	648	37	25.2	1240	1	VHWVEE	structural polypro
576	37	25.2	345	1	NBMS	apolipoprotein H p	649	37	25.2	1241	2	S26373	genome polyprotein
577	37	25.2	346	2	T20458	hypothetical prote	650	37	25.2	1242	2	S72350	structural polypro
578	37	25.2	362	2	A44083	meg protein - Mare	651	37	25.2	1242	2	A56605	structural polypro
579	37	25.2	365	2	S43417	RNA/DNA-binding pr	652	37	25.2	1287	2	A41685	SIL protein - huma
580	37	25.2	365	2	B70435	hypothetical prote	653	37	25.2	1313	2	T29027	hypothetical prote
581	37	25.2	369	2	S21471	genome polyprotein	654	37	25.2	1332	2	T23024	hypothetical prote
582	37	25.2	373	2	T36591	hypothetical prote	655	37	25.2	1344	3	T34188	myb-binding protei
583	37	25.2	375	1	TDHUMA	monocyte surface g	656	37	25.2	1375	3	T30813	plasminogen relate
584	37	25.2	375	2	B75220	replication factor	657	37	25.2	1376	1	VGJHJ2	E2 glycoprotein pr
585	37	25.2	376	2	T05420	cyclin delta-3 - A	658	37	25.2	1437	2	FQ1534	E2 glycoprotein pr
586	37	25.2	386	2	T22666	hypothetical prote	659	37	25.2	1437	2	F69680	DNA-directed DNA p
587	37	25.2	386	2	S52035	probable alcohol d	660	37	25.2	1443	2	S73446	DNA polymerase III
588	37	25.2	386	2	T05781	hypothetical prote	661	37	25.2	1468	2	S58250	DNA-directed DNA p
589	37	25.2	390	2	S52036	probable alcohol d	662	37	25.2	1506	2	T32909	hypothetical prote
590	37	25.2	395	2	D70441	oxido/reductase ir	663	37	25.2	1612	2	T30805	cutti protein - mo
591	37	25.2	396	1	A58938	surface protein rh	664	37	25.2	1616	2	T17884	S-layer protein -
592	37	25.2	405	2	G72655	probable aspartate	665	37	25.2	1616	2	G64242	cytadherence-acces
593	37	25.2	408	2	B75635	phosphoenolpyruvat	666	37	25.2	1651	2	T14160	transmembrane rece
594	37	25.2	410	2	S64451	hypothetical prote	667	37	25.2	1734	2	A54602	microtubule-associ
595	37	25.2	424	1	WFPGBA	inhibin beta-A cha	668	37	25.2	1921	2	T13827	kinesin-73 - fruit
596	37	25.2	424	1	S31440	inhibin beta-A cha	669	37	25.2	2111	2	T15390	hypothetical prote
597	37	25.2	424	1	B40905	inhibin beta-A cha	670	37	25.2	2126	2	E70522	probable polyketid
598	37	25.2	425	1	S50898	inhibin beta-A cha	671	37	25.2	2358	2	T39569	probable alpha-glu
599	37	25.2	425	1	T47072	inhibin beta-A cha	672	37	25.2	2364	2	A56577	microtubule-associ
600	37	25.2	426	1	B24248	inhibin beta-A cha	673	37	25.2	2464	1	QRMSPI	microtubule-associ
601	37	25.2	432	2	T48343	interleukin-11 rec	674	37	25.2	2531	2	S18188	notch protein homo
602	37	25.2	447	2	S66256	alpha-1,6-mannosyl	675	37	25.2	2605	2	T18552	saframycin Mxi syn
603	37	25.2	458	2	A48392	alpha 2C4 adrenoce	676	37	25.2	2871	2	A55567	fibrellin I - bovi
604	37	25.2	459	2	B44498	radial spoke prote	677	37	25.2	3006	2	T28625	variant-specific s
605	37	25.2	468	2	T17306	hypothetical prote	678	37	25.2	3164	1	WMBEH6	UL36 protein - hum
606	37	25.2	471	2	S08333	exo-alpha-sialidas	679	37	25.2	3605	2	S20901	titin - rabbit (fr
607	37	25.2	475	1	YWBO	cryptophan--tRNA l	680	37	25.2	10797	3	T30192	probable peptide s
608	37	25.2	485	2	S48650	catalase (EC 1.11.	681	36.5	24.8	56	2	D31445	ovomucoid, third d
609	37	25.2	492	2	S59733	tyrosine--tRNA lig	682	36.5	24.8	56	2	E31445	ovomucoid, third d
610	37	25.2	492	2	S50554	hypothetical prote	683	36.5	24.8	63	2	S08190	metallothionein 1
611	37	25.2	497	2	S11891	anthranilate synth	684	36.5	24.8	104	1	B41903	arsenical resistan
612	37	25.2	501	2	S22384	dihydrolipoamide d	685	36.5	24.8	104	1	A41902	arsenical resistan
613	37	25.2	516	2	C44479	collagen alpha 1(X	686	36.5	24.8	131	2	S78133	ribosomal protein

687	36.5	24.8	209	2	B70689	hypothetical prote	760	36	24.5	165	2	A69400	molybdopterin oxid
688	36.5	24.8	212	2	S43601	R07E5.4 protein (c	761	36	24.5	179	1	F69275	hypothetical prote
689	36.5	24.8	222	2	S77557	hypothetical prote	762	36	24.5	186	2	S08514	cytotoxic T-lympho
690	36.5	24.8	233	1	S13625	eosinophil major b	763	36	24.5	190	2	T20725	hypothetical prote
691	36.5	24.8	249	2	T24604	hypothetical prote	764	36	24.5	193	1	S73866	hypothetical prote
692	36.5	24.8	253	2	E75465	hypothetical prote	765	36	24.5	193	2	T15681	hypothetical prote
693	36.5	24.8	261	2	T40111	l4p-like ribosomal	766	36	24.5	194	1	R5PM24	ribosomal protein
694	36.5	24.8	267	2	S09673	kafirin precursor	767	36	24.5	199	2	T29943	hypothetical prote
695	36.5	24.8	299	2	T03792	proline-rich prote	768	36	24.5	200	2	I48615	gene KIS protein -
696	36.5	24.8	304	2	A33274	insulin-like growt	769	36	24.5	206	2	A65189	y19W protein - Esc
697	36.5	24.8	310	2	A60967	insulin-like growt	770	36	24.5	207	2	A75475	probable acetyltra
698	36.5	24.8	317	2	I46916	insulin-like growt	771	36	24.5	209	1	X0HYMC	methyated-DNA--pr
699	36.5	24.8	321	2	T15634	hypothetical prote	772	36	24.5	210	2	S73646	hypothetical prote
700	36.5	24.8	328	2	G71838	probable nadh oxid	773	36	24.5	215	2	S77007	hypothetical prote
701	36.5	24.8	360	2	F70859	hypothetical prote	774	36	24.5	219	2	T39945	ribosomal protein
702	36.5	24.8	365	2	A48135	transcription nega	775	36	24.5	223	2	I46696	CTLA-4 precursor -
703	36.5	24.8	425	2	S34449	transcription fact	776	36	24.5	223	2	A29063	cytotoxic T-lympho
704	36.5	24.8	429	2	A47305	translation initia	777	36	24.5	223	2	T09536	cytotoxic T-lympho
705	36.5	24.8	431	2	T12450	hypothetical prote	778	36	24.5	224	1	W0BERG	U11 protein - huma
706	36.5	24.8	436	2	S63398	probable membrane	779	36	24.5	224	2	S00959	hypothetical prote
707	36.5	24.8	437	1	A31752	transcription fact	780	36	24.5	232	1	K0PG	tissue kallikrein
708	36.5	24.8	437	2	S42111	transcription fact	781	36	24.5	243	2	C69316	molybdenum cofacto
709	36.5	24.8	459	2	T10526	cyclin B1c-11 - ye	782	36	24.5	244	2	C65016	hypothetical prote
710	36.5	24.8	463	2	T10015	hypothetical prote	783	36	24.5	253	1	G69109	conserved hypotet
711	36.5	24.8	473	2	E64120	exodeoxyribonuclea	784	36	24.5	253	2	JC4594	cell surface RAF-1
712	36.5	24.8	482	2	G70199	probable transcrip	785	36	24.5	258	1	B49904	probable regulator
713	36.5	24.8	510	2	S56233	probable membrane	786	36	24.5	259	1	C47069	acetylactate decar
714	36.5	24.8	541	2	T15299	hypothetical prote	787	36	24.5	260	2	T06326	malate dehydrogena
715	36.5	24.8	550	2	S75122	acetylactate synth	788	36	24.5	260	2	C70115	ribosomal protein
716	36.5	24.8	551	2	S66695	probable membrane	789	36	24.5	261	2	S57655	glutinin low molec
717	36.5	24.8	554	1	F0HUMP	macrophage colony-	790	36	24.5	267	2	G02764	cyto G1 interact
718	36.5	24.8	570	2	T08075	4-coumarate--CoA 1	791	36	24.5	276	2	S57656	glutinin low molec
719	36.5	24.8	580	2	T40939	probable Ca-calmod	792	36	24.5	277	2	C38277	hydroxyproline-ric
720	36.5	24.8	626	2	S37622	proto-oncogene - m	793	36	24.5	277	2	A46241	interferon respons
721	36.5	24.8	639	2	JC4881	polyvinyl-alcohol	794	36	24.5	281	1	ZBBE12	30.2K zinc-binding
722	36.5	24.8	643	1	S15623	E1 protein - human	795	36	24.5	285	2	I51412	hypothetical trans
723	36.5	24.8	643	2	S36499	E1 protein - human	796	36	24.5	286	2	F69801	epoxide hydrolase
724	36.5	24.8	702	2	S65544	CAMP-dependent pho	797	36	24.5	289	2	F72637	probable alcohol d
725	36.5	24.8	714	2	S65542	CAMP-dependent pho	798	36	24.5	290	2	C65053	hydrogenase isoenz
726	36.5	24.8	736	2	T25447	hypothetical prote	799	36	24.5	302	2	A70711	hypothetical prote
727	36.5	24.8	754	2	T25551	hypothetical prote	800	36	24.5	307	2	S10015	alpha/beta-gliadin
728	36.5	24.8	777	2	S65543	3',5'-cyclic-nucle	801	36	24.5	309	1	S60157	RING finger protei
729	36.5	24.8	832	2	T31878	hypothetical prote	802	36	24.5	309	1	A57235	RING finger protei
730	36.5	24.8	846	2	T27282	hypothetical prote	803	36	24.5	309	2	S60120	RING finger protei
731	36.5	24.8	865	2	B69074	probable formate d	804	36	24.5	309	2	I53384	4-1BB ligand - mou
732	36.5	24.8	865	2	T32574	hypothetical prote	805	36	24.5	313	2	S07524	alpha/beta-gliadin
733	36.5	24.8	1019	2	T13039	tyrosine kinase re	806	36	24.5	316	2	S41033	hypothetical prote
734	36.5	24.8	1057	2	A42109	glycine dehydrogen	807	36	24.5	323	1	TVWVF6	protein kinase (EC
735	36.5	24.8	1059	2	T22545	hypothetical prote	808	36	24.5	329	2	T32760	hepatocellular car
736	36.5	24.8	1075	2	D70568	hypothetical prote	809	36	24.5	333	1	A41099	hemoglobin precurs
737	36.5	24.8	1081	1	A42399	isoleucine--tRNA 1	810	36	24.5	338	2	C75459	probable endonucle
738	36.5	24.8	1547	2	T28657	blackjack protein,	811	36	24.5	340	2	S32584	structural protein
739	36.5	24.8	1610	2	T11681	hypothetical prote	812	36	24.5	341	2	B42465	regulatory protein
740	36.5	24.8	2470	2	S57085	1-phosphatidylinos	813	36	24.5	342	2	PC4211	hepatocellular car
741	36.5	24.8	2523	2	F70846	probable pPE prote	814	36	24.5	351	2	JC5121	oligopeptide perme
742	36.5	24.8	9376	2	T14593	syringomycin synth	815	36	24.5	354	2	T32246	hypothetical prote
743	36	24.5	25	2	T01689	ATPase-beta chain	816	36	24.5	356	2	S01992	glutinin low molec
744	36	24.5	43	2	S07621	avenin gamma-3 - s	817	36	24.5	356	2	T09573	replication factor
745	36	24.5	52	1	A33173	rubredoxin - Clost	818	36	24.5	367	2	T24298	hypothetical prote
746	36	24.5	54	2	I39520	rubredoxin - Acine	819	36	24.5	368	2	A54430	hypoxic function t
747	36	24.5	60	2	T11853	ribosomal protein	820	36	24.5	369	2	A60041	Ca2+/calmodulin-d
748	36	24.5	83	2	S77417	prochlorophyllid	821	36	24.5	369	2	JQ2278	hydroxymethylbilan
749	36	24.5	93	1	W7WL	E7 protein - human	822	36	24.5	374	2	T25809	hypothetical prote
750	36	24.5	111	2	Q00098	hypothetical l2K p	823	36	24.5	377	2	S28184	Ca2+/calmodulin-d
751	36	24.5	114	2	S77591	ferredoxin [2Fe-2S	824	36	24.5	380	2	JC1451	Ca2+/calmodulin-d
752	36	24.5	120	2	S43030	hypothetical prote	825	36	24.5	384	2	T40502	hypothetical prote
753	36	24.5	128	2	JQ1002	keratin, claw - ch	826	36	24.5	385	2	B72022	cell division prot
754	36	24.5	131	2	S64940	probable membrane	827	36	24.5	385	2	G71474	probable cell divi
755	36	24.5	138	1	TTFGB	thyrotropin beta c	828	36	24.5	389	2	S25854	ribonucleoside-dip
756	36	24.5	138	2	S75195	ferric uptake regu	829	36	24.5	392	2	D70656	hypothetical prote
757	36	24.5	138	2	JQ1543	mobB protein - Sal	830	36	24.5	395	2	T09373	hypothetical prote
758	36	24.5	145	2	H69954	transcription regu	831	36	24.5	400	2	A69270	molybdenum cofacto
759	36	24.5	165	1	D69285	conserved hypotet	832	36	24.5	404	2	A54871	Gal beta-1, 3GalNA

C:Keywords: glycoprotein

Query Match 61.9%; Score 91; DB 2; Length 120;
Best Local Similarity 64.3%; Pred. No. 7.4e-06;
Matches 18; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

QY 2 pd----vqdcpeclqenpfsspgapi 25
|| :| ||||| || |||:|||||

Db 26 PDGEFLMQGCEKLGNGRFFSKPGAPI 53

RESULT 5

S74086

follicotropin alpha chain - ostrich

C:Species: Struthio camelus (ostrich)

C>Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 17-Apr-1998

C:Accession: S74086

R:Koide, Y.; Papkoff, H.; Kawauchi, H.

Eur. J. Biochem. 240, 262-267, 1996

A:Title: Complete amino acid sequences of follitropin and lutropin in the ostrich, *Struthio camelus*

A:Reference number: S74084; MUID:97025333

A:Accession: S74086

A:Molecule type: protein

A:Residues: 1-96 <KOI>

A:Experimental source: pituitary glands

C:Superfamily: glycoprotein hormones alpha chain

C:Keywords: glycoprotein; heterodimer; hormone; pituitary

F:11-35,14-64,32-86,36-88,63-91/Disulfide bonds: #status predicted

F:56,82/Binding.site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 61.2%; Score 90; DB 2; Length 96;

Matches 17; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY 2 pd----vqdcpeclqenpfsspgapi 25
|| :| ||||| || |||:|||||

Db 2 PDGEFLMQGCEKLGNGRFFSKPGAPV 29

RESULT 6

TBEOA

glycoprotein hormones alpha chain precursor - bovine

N:Alternate names: chorionadotropin alpha chain; follicle-stimulating hormone alpha chain; lutropin alpha chain; thyrotropin alpha chain

C:Species: Bos primigenius taurus (cattle)

C>Date: 24-Apr-1984 #sequence_revision 30-Sep-1987 #text_change 23-Aug-1996

C:Accession: A93489; A05132; B94673; A92099; A01483; A05270

R:Goodwin, R.G.; Moncman, C.L.; Rottman, F.M.; Nilson, J.H.

Nucleic Acids Res. 11, 6873-6882, 1983

A:Title: Characterization and nucleotide sequence of the gene for the common alpha subunit of the bovine pituitary gonadotropin

A:Reference number: A93489; MUID:84041490

A:Accession: A93489

A:Molecule type: DNA

A:Residues: 1-120 <GOO>

A:Cross-references: GB:X00002

R:Nilson, J.H.; Thomson, A.R.; Cserbak, M.T.; Moncman, C.L.; Woychik, R.P.

J. Biol. Chem. 258, 4679-4682, 1983

A:Title: Nucleotide sequence of a cDNA for the common alpha subunit of the bovine pituitary gonadotropin

A:Reference number: A05132; MUID:83161058

A:Accession: A05132

A:Molecule type: mRNA

A:Residues: 8-120 <NIL>

A:Cross-references: GB:J00009

R:Liao, T.H.; Pierce, J.G.

J. Biol. Chem. 246, 850-865, 1971

A:Title: The primary structure of bovine thyrotropin.

A:Reference number: A94673; MUID:71111428

A:Accession: B94673

A:Molecule type: protein

A:Residues: 25-120 <LIA>

R:Pierce, J.G.; Liao, T.H.; Carlsen, R.B.; Reimo, T.

J. Biol. Chem. 246, 866-872, 1971

A:Title: Comparisons between the alpha chain of bovine thyrotropin and the CI chain of bovine thyrotropin

A:Reference number: A92099; MUID:7111429

A:Accession: A92099

A:Molecule type: protein

A:Residues: 25-111,'SC',114-120 <PIE>

R:Corneil, J.S.; Pierce, J.G.

J. Biol. Chem. 249, 4166-4174, 1974

A:Title: Studies on the disulfide bonds of glycoprotein hormones. Locations in the alpha chain of bovine thyrotropin

A:Reference number: A92152; MUID:74307752

A:Contents: annotation; preliminary disulfide bonds

A:Note: disulfide bonds shown follow those determined crystallographically for human thyrotropin

C:Comment: Glycoprotein hormones are dimeric compounds with a common alpha chain and

C:Genetics:

A:Introns: 34/1; 95/3

C:Superfamily: glycoprotein hormones alpha chain

C:Keywords: glycoprotein; heterodimer; hormone; pituitary

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-120/Product: glycoprotein hormones alpha chain, long form #status experimental <F>

F:27-120/Product: glycoprotein hormones alpha chain, short form #status experimental <F>

F:35-59,38-88,56-110,60-112/Disulfide bonds: #status predicted

F:80,106/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:87-115/Disulfide bonds: #status experimental

Query Match

Best Local Similarity 57.8%; Score 85; DB 1; Length 120;

Matches 16; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

QY 2 pd----vqdcpeclqenpfsspgapi 25
|| :| ||||| || |||:|||||

Db 26 PDGEFLMQGCEKLGNGRFFSKPGAPI 53

RESULT 7

UTSHA

glycoprotein hormones alpha chain precursor - sheep

N:Alternate names: chorionadotropin alpha chain; follicle-stimulating hormone alpha chain; lutropin alpha chain; thyrotropin alpha chain

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C>Date: 28-Aug-1985 #sequence_revision 02-May-1994 #text_change 18-Jun-1999

C:Accession: S06935; A92109; A90312; A61098; S13200; A01484

R:Bello, P.A.; Mountford, P.S.; Brandon, M.R.; Adams, T.E.

Nucleic Acids Res. 17, 10494, 1989

A:Title: Cloning and DNA sequence analysis of the cDNA for the common alpha-subunit of the sheep pituitary gonadotropin

A:Reference number: S06935; MUID:90098887

A:Accession: S06935

A:Molecule type: mRNA

A:Residues: 1-120 <BEL>

A:Cross-references: EMBL:X16977; NID:gl365; PIDN:CAA34848.1; PID:gl366

R:Liu, W.K.; Nahm, H.S.; Sweeney, C.M.; Lamkin, W.M.; Baker, H.N.; Ward, D.N.

J. Biol. Chem. 247, 4351-4364, 1972

A:Title: The primary structure of ovine luteinizing hormone. I. The amino acid sequence of the primary structure of ovine luteinizing hormone

A:Reference number: A92109; MUID:72211144

A:Contents: lutropin

A:Accession: A92109

A:Molecule type: protein

A:Residues: 25-26,'N',28,'Q',30-36,'Q',38-120 <LIU>

A:Note: amides were not determined

R:Sairam, M.R.

Biochem. J. 197, 535-539, 1981

A:Title: Primary structure of the ovine pituitary follitropin alpha-subunit.

A:Reference number: A90312; MUID:82113052

A:Contents: follitropin

A:Accession: A90312

A:Molecule type: protein

A:Residues: 25-26,'N',28,'Q',30-36,'Q',38-120 <LI2>

A:Note: amides were not determined

R:Namura, K.; Tsunawake, S.; Ohmura, K.; Sakiyama, F.; Shizume, K.

Endocrinology 123, 700-712, 1988

A:Title: Renotropic activity in ovine luteinizing hormone isoform(s).

A:Reference number: A61098; MUID:88283534

A:Accession: A61098
A:Molecule type: protein
A:Residues: 25-41 <NOM>
A:Note: this form was designated form alpha-3; forms alpha-1 and alpha-2 each lack several amino acids
R:Chung, D.; Sairam, M.R.; Li, C.H.
Arch. Biochem. Biophys. 159, 678-682, 1973
A:Title: The primary structure of ovine interstitial cell-stimulating hormone. III: Disulfide bonds
A:Reference number: A90057
A:Contents: annotation; lutropin, preliminary disulfide bonds
A:Note: disulfide bonds shown follow those determined crystallographically for human glycoprotein hormone
R:Weishaar, G.; Hiyama, J.; Renwick, A.G.C.
Eur. J. Biochem. 192, 741-751, 1990
A:Title: Site-specific N-glycosylation of ovine lutropin. Structural analysis by one-dimensional peptide mapping
A:Reference number: S13200; MUID:91006170
A:Accession: S13200
A>Status: preliminary
A:Molecule type: protein
A:Residues: 80-91;104-119 <WEI>
C:Superfamily: glycoprotein hormones alpha chain
C:Keywords: glycoprotein; heterodimer; hormone; pituitary
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-120/Product: glycoprotein hormones alpha chain #status experimental <MAT>
F:35-59,38-88,56-110,60-112,87-115/Disulfide bonds: #status predicted
F:80,106/Binding site: carbohydrate (Asn) (covalent) #status experimental
Query Match 57.8%; Score 85; DB 1; Length 120;
Best Local Similarity 57.1%; Pred. No. 5.1e-05;
Matches 16; Conservative 4; Mismatches 4; Indels 4; Gaps 1;
Qy 2 pd----vqdcpectlqenpfsgqapi 25
Db 26 PDGFTWQGCPECKLKENKFSKLGAPI 53

RESULT 8
TTTFA
glycoprotein hormones alpha chain precursor - rat
N:Alternate names: chorionadotropin alpha chain; follicle-stimulating hormone alpha chain; lutropin alpha chain; thyrotropin alpha chain
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 18-Jun-1999
C:Accession: J0408; S27385; S27386; I83050
R:Burnside, J.; Buckland, P.R.; Chin, W.W.
Gene 70, 67-74, 1988
A:Title: Isolation and characterization of the gene encoding the alpha-subunit of the rat chorionadotropin
A:Reference number: J0408; MUID:89196918
A:Accession: J0408
A:Molecule type: DNA
A:Residues: 1-120 <BOR>
A:Cross-references: GB:M22830
R:Godine, J.E.; Chin, W.W.; Habener, J.F.
J. Biol. Chem. 257, 8368-8371, 1982
A:Title: Alpha subunit of rat pituitary glycoprotein hormones. Primary structure of the gene
A:Reference number: S27385; MUID:82214055
A:Accession: S27385
A:Molecule type: mRNA
A:Residues: 1-120 <GOD>
A:Cross-references: EMBL:V01252; NID:g56579; PIDN:CAA24565.1; PID:g56580
A:Accession: S27386
A:Molecule type: mRNA
A:Residues: 1-2, 'Y', '4', 'K', '6-17', 'F', '19-29', 'F', '31-91', 'A', '93-108', 'E', '110-120 <GOD>
A:Cross-references: EMBL:V01253; NID:g56582; PIDN:CAA24566.1; PID:g56583
R:Kato, Y.; Ezashi, T.; Hirai, T.; Kato, T.
Zool. Sci. 7, 879-887, 1990
A:Title: Strain difference in Nucleotide Sequences of Rat Glycoprotein Hormone Subunit c
A:Reference number: I60104
A:Accession: I83050
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-120 <RES>
A:Cross-references: GB:D00575; NID:g220841; PIDN:BAA00453.1; PID:g220842
C:Genetics:

A:Introns: 34/1; 95/3
C:Superfamily: glycoprotein hormones alpha chain
C:Keywords: glycoprotein; heterodimer; hormone; pituitary
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-120/Product: glycoprotein hormones alpha chain #status predicted <MAT>
F:35-59,38-88,56-110,60-112,87-115/Disulfide bonds: #status predicted
F:80,106/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 57.1%; Score 84; DB 1; Length 120;
Best Local Similarity 57.1%; Pred. No. 7.1e-05;
Matches 16; Conservative 4; Mismatches 4; Indels 4; Gaps 1;
Qy 2 pd----vqdcpectlqenpfsgqapi 25
Db 26 PDGDLIIQGCPECKLKENKFSKLGAPI 53

RESULT 9
TTMSA
glycoprotein hormones alpha chain precursor - mouse
N:Alternate names: chorionadotropin alpha chain; follicle-stimulating hormone alpha chain; lutropin alpha chain; thyrotropin alpha chain
C:Species: Mus musculus (house mouse)
C:Date: 15-Oct-1982 #sequence_revision 15-Oct-1982 #text_change 18-Jun-1999
C:Accession: A31598; A01482
R:Gordon, D.F.; Wood, W.M.; Ridgway, E.C.
DNA 7, 679-690, 1988
A:Title: Organization and nucleotide sequence of the mouse alpha-subunit gene of the chorionadotropin
A:Reference number: A31598; MUID:89170115
A:Accession: A31598
A:Molecule type: DNA
A:Residues: 1-120 <GOR>
A:Cross-references: GB:M22992; NID:g340780; PIDN:AAA99228.1; PID:g575520
R:Chin, W.W.; Kronenberg, H.M.; Dee, P.C.; Maloof, F.; Habener, J.F.
Proc. Natl. Acad. Sci. U.S.A. 78, 5329-5333, 1981
A:Title: Nucleotide sequence of the mRNA encoding the pre-alpha-subunit of mouse thyrotropin
A:Reference number: A01482; MUID:82060239
A:Accession: A01482
A:Molecule type: mRNA
A:Residues: 1-120 <CHI>
A:Cross-references: GB:V00852; NID:g54798; PIDN:AAA96700.1; PID:g54799; GB:J00643; GB:J00644
C:Genetics:
A:Introns: 34/1; 95/3
C:Superfamily: glycoprotein hormones alpha chain
C:Keywords: glycoprotein
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-120/Product: glycoprotein hormones alpha chain #status predicted <MAT>
F:35-59,38-88,56-110,60-112,87-115/Disulfide bonds: #status predicted
F:80,106/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 57.1%; Score 84; DB 1; Length 120;
Best Local Similarity 57.1%; Pred. No. 7.1e-05;
Matches 16; Conservative 4; Mismatches 4; Indels 4; Gaps 1;
Qy 2 pd----vqdcpectlqenpfsgqapi 25
Db 26 PDGDLIIQGCPECKLKENKFSKLGAPI 53

RESULT 10
A05096
glycoprotein hormones alpha chain - rabbit
N:Alternate names: lutropin alpha chain
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 07-Feb-1997
C:Accession: A05096
R:Glenn, S.D.; Nahn, H.S.; Ward, D.N.
J. Protein Chem. 3, 143-156, 1984
A:Title: The amino acid sequence of the rabbit glycoprotein hormone alpha subunit.
A:Reference number: A05096
A:Accession: A05096

A:Molecule type: protein
A:Residues: 1-96 <GLE>
C:Superfamily: glycoprotein hormones alpha chain
C:Keywords: glycoprotein

Query Match 55.8%; Score 82; DB 2; Length 96;
Best Local Similarity 57.1%; Pred. No. 0.00011;
Matches 16; Conservative 4; Mismatches 4; Indels 1;

QY 2 pd----vqdcpectlqenpfssqgapi 25
|| : ||||| : || : |||||
Db 2 PDGEFAMQGPCPECKLKKNYFSKLGAPI 29

RESULT 11

UTPGA

glycoprotein hormones alpha chain precursor - pig
N:Alternate names: choriongonadotropin alpha chain; follicle-stimulating hormone alpha chain
N: lutropin alpha chain; thyrotropin alpha chain

C:Species: Sus scrofa domestica (domestic pig)
C:Date: 24-Apr-1984 #sequence_revision 20-Jan-1995 #text_change 18-Jun-1999
C:Accession: A30339; A91086; A91213; B60584; I46618; A01485

R:Hirai, T.; Takikawa, H.; Kato, Y.
Mol. Cell. Endocrinol. 63, 209-217, 1989
A:Title: Molecular cloning of cDNAs for precursors of porcine pituitary glycoprotein hormones

A:Reference number: A30339; MUID:89325834
A:Accession: A30339

A>Status: not compared with conceptual translation
A:Molecule type: mRNA

A:Residues: 1-120 <HIR>
A>Note: authors note that the mature amino-terminal hexapeptide was probably lost during

R:Closset, J.; Maguin-Rogister, G.; Hennen, G.
Endocrinol. Exp. 8, 164, 1974

A:Title: The amino acid sequence of porcine thyrotropin (TSH) with reference to the molecule
A:Reference number: A91086

A:Contents: thyrotropin
A:Accession: A91086

A:Molecule type: protein
A:Residues: 31-119 <CLO>

R:Maguin-Rogister, G.; Combarnous, Y.; Hennen, G.
Eur. J. Biochem. 39, 255-263, 1973

A:Title: The primary structure of the porcine luteinizing-hormone alpha-subunit.
A:Reference number: A91213; MUID:74075725

A:Contents: lutropin
A:Accession: A91213

A:Molecule type: protein
A:Residues: 31-119 <MAG>

R:Combarnous, Y.; Hennen, G.
Biochem. Soc. Trans. 2, 915-917, 1974

A:Title: The disulphide bridges of porcine luteinizing hormone alpha subunit.
A:Reference number: A90350; MUID:75093922

A:Contents: annotation; lutropin, preliminary disulfide bonds
A>Note: disulfide bonds shown follow those determined crystallographically for human gly

R:Nomura, K.; Ohmura, K.; Nakamura, Y.; Horiba, N.; Shirakura, Y.; Sato, Y.; Ujihara, M.
Endocrinology 124, 712-719, 1989

A:Title: Porcine luteinizing hormone isoform(s): relationship between their molecular st
A:Reference number: A60584; MUID:89107050

A:Accession: B60584
A:Molecule type: protein

A:Residues: 25-34 <NOM>
R:Kato, Y.; Ezashi, T.; Hirai, T.; Kato, T.

J. Mol. Endocrinol. 7, 27-34, 1991
A:Title: The gene for the common alpha subunit of porcine pituitary glycoprotein hormones

A:Reference number: I46618; MUID:91369449
A:Accession: I46618

A>Status: translated from GB/EMBL/DBDJ
A:Residues: 1-120 <KAT>

A:Cross-references: GB:D00768; NID:9217697; PIDN:BAA00664.1; PID:9217699
C:Genetics:

A:Introns: 33/3; 95/3
C:Superfamily: glycoprotein hormones alpha chain

C:Keywords: glycoprotein; heterodimer; hormone; pituitary
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-120/Product: glycoprotein hormones alpha chain #status predicted <MAT>
F:35-59,56-110,60-112,87-115/Disulfide bonds: #status predicted
F:38-88/Disulfide bonds: #status experimental
F:80,106/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 55.8%; Score 82; DB 1; Length 120;
Best Local Similarity 57.1%; Pred. No. 0.00014;
Matches 16; Conservative 4; Mismatches 4; Indels 1;

QY 2 pd----vqdcpectlqenpfssqgapi 25
|| : ||||| : || : |||||
Db 26 PDGEFTMQGPCPECKLKKNYFSKLGAPI 53

RESULT 12

S53062

glycoprotein hormones alpha chain precursor - donkey

C:Species: Equus asinus (donkey)
C:Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 16-Jul-1999
C:Accession: S53062

R:Chopineau, M.; Stewart, F.
submitted to the EMBL Data Library, March 1995
A:Description: Cloning and analysis of the cDNA for the common alpha subunit of the d

A:Reference number: S53062
A:Accession: S53062

A>Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-120 <CHO>
A:Cross-references: EMBL:X85170; NID:9732677; PIDN:CAA59454.1; PID:9732678

C:Superfamily: glycoprotein hormones alpha chain
C:Keywords: glycoprotein

Query Match 54.4%; Score 80; DB 2; Length 120;
Best Local Similarity 53.8%; Pred. No. 0.00026;
Matches 15; Conservative 4; Mismatches 5; Indels 1;

QY 2 pd----vqdcpectlqenpfssqgapi 25
|| : ||||| : || : |||||
Db 26 PDGEFTQDCPECKLKKNYFSKLGAPI 53

RESULT 13

TTHOA

glycoprotein hormones alpha chain - horse

N:Alternate names: choriongonadotropin alpha chain; follicle-stimulating hormone alpha
N: lutropin alpha chain; thyrotropin alpha chain
C:Species: Equus caballus (domestic horse)

C:Date: 31-May-1979 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C:Accession: A28576; A01487; A01486

R:Stewart, F.; Thomson, J.A.; Leigh, S.E.A.; Warwick, J.M.
J. Endocrinol. 115, 341-346, 1987

A:Title: Nucleotide (cDNA) sequence encoding the horse gonadotrophin alpha-subunit.
A:Reference number: A28576; MUID:88140711

A:Accession: A28576
A:Molecule type: mRNA

A:Residues: 4-96 <STE>
A:Cross-references: GB:M27462; NID:9602447; PIDN:AAA57252.1; PID:9602448

R:Moore Jr., W.T.; Ward, D.N.; Burtleigh, B.D.
Fed. Proc. 38, 462, 1979

A:Title: Saccharide products of cultured human mammary cells.
A:Reference number: A01487

A:Accession: A01487
A:Molecule type: protein

A:Residues: 1-8,'EB',11-59,'Z',61-80,'Z',82-84,'Z',86-96 <MOO>
R:Rathnam, P.; Fujiki, Y.; Landefeld, T.D.; Saxena, B.B.

J. Biol. Chem. 253, 5355-5362, 1978
A:Title: Isolation and amino acid sequence of the alpha-subunit of follicle-stimulat

A:Reference number: A01486; MUID:78218212
A:Accession: A01486

C: Superfamily: glycoprotein hormones alpha chain
 F: 1-33/domain: signal sequence #status predicted <SIG>
 F: 24-118/Product: glycoprotein hormones alpha-1 chain #status predicted <MAT>
 F: 34-58, 37-87, 95-108, 99-110, 86-113/disulfide bonds: #status predicted
 F: 79, 104/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	Score	DB	Length
Best Local Similarity	49.7%;	73;	118;
	63.2%;		
Pred. NO.	0.0024;		

Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 7 cpecltqlenpfssqpgapi 25

|||||:|||||

Db 34 CEECKLKNNIFSKPGAPV 52

RESULT 18

S16762

gonadotropin alpha chain - grass carp

C:Species: Ctenopharyngodon idella (grass carp)

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 16-Jul-1999

C:Accession: S16762

R:Chang, Y.S.; Huang, F.L.; Lo, T.B.

Submitted to the EMBL data library, July 1991

A:Description: The cDNA cloning and primary structures of grass carp gonadotropin subunit

A:Accession: S16762

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-118 <CHA>

A:Cross-references: EMBL:X61050; NID:g62706; PIDN:CAA43384.1; PID:g62707

C:Superfamily: glycoprotein hormones alpha chain

Query Match

Best Local Similarity 49.7%; Score 73; DB 2; Length 118;

Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 7 cpecltqlenpfssqpgapi 25

|||||:|||||

Db 34 CEECKLKNNIFSKPGAPV 52

RESULT 19

A60626

glycoprotein hormones alpha chain precursor - silver carp

N:Alternate names: gonadotropin alpha chain

C:Species: Hypophthalmichthys molitrix (silver carp)

C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 07-May-1999

C:Accession: A60626

R:Chang, Y.S.; Huang, C.J.; Huang, F.L.; Liu, C.S.; Lo, T.B.

Gen. Comp. Endocrinol. 78, 23-33, 1990

A:Title: Purification, characterization, and molecular cloning of gonadotropin subunits

A:Reference number: A60626; MUID:90236229

A:Accession: A60626

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-118 <CHA>

A:Note: parts of this sequence, including the amino and carboxyl ends of the mature protein

C:Superfamily: glycoprotein hormones alpha chain

C:Keywords: glycoprotein; hormone; pituitary

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-118/Product: glycoprotein hormones alpha chain, long form #status experimental <MAT

F:26-118/Product: glycoprotein hormones alpha chain, short form #status experimental <MA

F:79,104/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 49.7%; Score 73; DB 2; Length 118;

Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 7 cpecltqlenpfssqpgapi 25

|||||:|||||

Db 34 CEECKLKNNIFSKPGAPV 52

RESULT 20

A40554

glycoprotein hormones alpha-2 chain precursor - common carp

N:Alternate names: gonadotropin alpha chain type II

C:Species: Cyprinus carpio (common carp)

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999

C:Accession: A40554

R:Chang, Y.S.; Huang, C.J.; Huang, F.L.; Lo, T.B.

Int. J. Pept. Protein Res. 32, 556-564, 1988

A:Title: Primary structures of carp gonadotropin subunits deduced from cDNA nucleotide

A:Reference number: JK0024; MUID:89233593

A:Accession: A40554

A:Molecule type: mRNA

A:Residues: 1-118 <CHA>

A:Cross-references: GB:M37380; NID:g213055; PIDN:AAA49210.1; PID:g213056

A:Note: the authors translated the codon GCU for residue 7 as Arg

C:Superfamily: glycoprotein hormones alpha chain

C:Keywords: glycoprotein; pituitary

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-95/Product: glycoprotein hormones alpha-2 chain #status predicted <MAT>

F:79,104/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 49.7%; Score 73; DB 2; Length 118;

Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 7 cpecltqlenpfssqpgapi 25

|||||:|||||

Db 34 CEECKLKNNIFSKPGAPV 52

RESULT 21

I51229

gonadotropin alpha 1 subunit - cherry salmon

C:Species: Oncorhynchus masou (cherry salmon)

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999

C:Accession: I51229

R:Gen, K.; Maruyama, O.; Kato, T.; Tomizawa, K.; Wakabayashi, K.; Kato, Y.

J. Mol. Endocrinol. 11, 265-273, 1993

A:Title: Molecular cloning of cDNAs encoding two types of gonadotropin alpha subunit

ubunits.

A:Reference number: I51229; MUID:94197892

A:Accession: I51229

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-119 <GEN>

A:Cross-references: GB:S69273; NID:g546257; PIDN:RAA30421.1; PID:g546258

C:Superfamily: glycoprotein hormones alpha chain

Query Match

Best Local Similarity 49.7%; Score 73; DB 2; Length 119;

Matches 12; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 7 cpecltqlenpfssqpgapi 25

|||||:|||||

Db 35 CEECKLKNNIFSNPGAPV 53

RESULT 22

S07091

glycoprotein hormones alpha chain - daggertooth pike conger

N:Alternate names: gonadotropin alpha chain

C:Species: Muraenesox cinereus (daggertooth pike conger)

C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 07-Feb-1997

C:Accession: S07091

R:Liu, C.S.; Huang, F.L.; Chang, Y.S.; Lo, T.B.

Eur. J. Biochem. 186, 105-114, 1989

A:Title: Pike eel (Muraenesox cinereus) gonadotropin. Amino acid sequences of both al

A:Reference number: S07091; MUID:90092087

A:Accession: S07091

A:Molecule type: protein

A:Residues: 1-93 <LIU>

C:Superfamily: glycoprotein hormones alpha chain

C:Keywords: glycoprotein

Query Match

Best Local Similarity 49.0%; Score 72; DB 2; Length 93;


```

Query Match      34.7%; Score 51; DB 2; Length 533;
Best Local Similarity 44.4%; Pred. No. 13;
Matches      8; Conservative      3; Mismatches      7; Indels      0; Gaps      0;

QY      8 pcvqgqfpggapi 25
      ||| | : | | | :
Db      173 PECDLQGSWFYDPGGPL 190

RESULT 28
FNBO
fibronectin - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Dec-1988 #sequence-revision 31-Dec-1988 #text-change 07-May-1999
C:Accession: A26452; B21165; A23292
R:Skorstengaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.E.; Magnusson, S.
Eur. J. Biochem. 161, 441-453, 1986
A:Title: Complete primary structure of bovine plasma fibronectin.
A:Reference number: A26452; MUID:87054047
A:Accession: A26452
A:Molecule type: protein
A:Residues: 1-2265 <SKO>
R:Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectins
A:Reference number: A21165; MUID:83221567
A:Accession: B21165
A:Molecule type: mRNA
A:Residues: 2170-2265 <KOR>
A:Cross-references: GB:K00800
R:Petersen, T.E.; Thogersen, H.C.; Skorstengaard, K.; Sahl, P.; Sott
Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983
A:Title: Partial primary structure of bovine plasma fibronectin: three types of internal
A:Reference number: A23292; MUID:83117805
A:Accession: A23292
A:Molecule type: protein
A:Residues: 1-16, 'C', 18-20, 'S', 22-432; 447-463; 1367-1517; 1567-1673; 2062-2176, 'N', 2178-226
C:Comment: Cys-1201 and Cys-2015 have free sulfhydryl groups.
C:Comment: The plasma fibronectin molecule consists of two chains, which are connected b
C:Comment: Fibronectins bind cell surfaces and various compounds including collagen, fib
aling, and maintenance of cell shape.
C:Comment: Plasma fibronectin is synthesized by hepatocytes.
C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C:Keywords: acute phase; alternative splicing; collagen binding; duplication; extracellu
F:21-241/Domain: fibrin and heparin binding <FBR>
F:21-56/Domain: fibronectin type I repeat homology <1F1>
F:66-104/Domain: fibronectin type I repeat homology <1F2>
F:110-148/Domain: fibronectin type I repeat homology <1F3>
F:155-194/Domain: fibronectin type I repeat homology <1F4>
F:200-239/Domain: fibronectin type I repeat homology <1F5>
F:277-577/Domain: collagen binding <CBR>
F:277-311/Domain: fibronectin type I repeat homology <1F6>
F:329-370/Domain: fibronectin type II repeat homology <2F1>
F:389-430/Domain: fibronectin type II repeat homology <2F2>
F:439-477/Domain: fibronectin type I repeat homology <1F7>
F:487-524/Domain: fibronectin type I repeat homology <1F8>
F:530-568/Domain: fibronectin type I repeat homology <1F9>
F:578-661/Domain: fibronectin type III repeat homology <FN3A>
F:688-770/Domain: fibronectin type III repeat homology <FN3B>
F:779-860/Domain: fibronectin type III repeat homology <FN3C>
F:875-957/Domain: fibronectin type III repeat homology <FN3D>
F:965-1046/Domain: fibronectin type III repeat homology <FN3E>
F:1055-1134/Domain: fibronectin type III repeat homology <FN3F>
F:1142-1227/Domain: fibronectin type III repeat homology <FN3G>
F:1235-1318/Domain: fibronectin type III repeat homology <FN3H>
F:1326-1404/Domain: fibronectin type III repeat homology <GN3I>
F:1410-1517/Domain: cell attachment <CAD>
F:1416-1502/Domain: fibronectin type III repeat homology <FN3J>
F:1493-1495/Region: cell attachment (R-G-D) motif
F:1510-1592/Domain: fibronectin type III repeat homology <FN3K>
F:1600-1682/Domain: heparin binding <HB2>
F:1692-1773/Domain: fibronectin type III repeat homology <FN3L>
F:1692-1773/Domain: fibronectin type III repeat homology <FN3M>

F:1781-1863/Domain: fibronectin type III repeat homology <FN3N>
F:1970-1972/Region: cell attachment (R-G-D) motif
F:1982-2062/Domain: fibronectin type III repeat homology <FN3O>
F:1985-2216/Domain: fibrin binding <FB2>
F:2085-2124/Domain: fibronectin type I repeat homology <1F10>
F:2130-2167/Domain: fibronectin type I repeat homology <1F11>
F:2174-2209/Domain: fibronectin type I repeat homology <1F12>
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:3/Cross-link: isopeptide (Gln) (interchain to fibrin) #status experimental
F:21-47, 45-56, 66-94, 92-104, 110-138, 136-148, 155-184, 182-194, 200-229, 227-239, 277-304, 30
7, 2155-2167, 2174-2200, 2198-2209/Disulfide bonds: #status predicted
F:399, 497, 511, 846, 976, 1213, 1987/Binding site: carbohydate (Asn) (covalent) #status e
F:1205, 1692/Binding site: carbohydate (Asn) (covalent) #status absent
F:1943, 1944/Binding site: carbohydate (Thr) (covalent) #status experimental
F:2246/Disulfide bonds: interchain (to 2250) #status predicted
F:2250/Disulfide bonds: interchain (to 2246) #status predicted
F:2263/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match      34.4%; Score 50.5; DB 1; Length 2265;
Best Local Similarity 42.9%; Pred. No. 62;
Matches      9; Conservative      5; Mismatches      6; Indels      1; Gaps      1;

QY      2 pdvqdcpectlqnpffsqg 22
      |:: | | | | | | | | |
Db      1887 PEILDVPS-IVQKTPFITNPG 1906

RESULT 29
S56594
hypothetical 30.1K protein (dnaT-hold intergenic region) - Escherichia coli
N:Alternate names: hypothetical protein f262b
C:Species: Escherichia coli
C:Date: 28-Oct-1995 #sequence-revision 03-Nov-1995 #text-change 08-Oct-1999
C:Accession: S56594; A65252
R:Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res. 23, 2105-2119, 1995
A:Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region from
A:Reference number: S56314; MUID:95334362
A:Accession: S56594
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-262 <BUR>
A:Cross-references: EMBL:U14003; NID:gl263172; PIDN:AAA97266.1; PID:g537210
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: A65252
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-262 <BLAT>
A:Cross-references: GB:AE000507; GB:U00096; NID:g2367380; PIDN:RAC77323.1; PID:g17908
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: yjiS
C:Superfamily: Escherichia coli hypothetical 30.1K protein (dnaT-hold intergenic regi

Query Match      34.0%; Score 50; DB 2; Length 262;
Best Local Similarity 66.7%; Pred. No. 8.7;
Matches      8; Conservative      2; Mismatches      2; Indels      0; Gaps      0;

QY      2 pdvqdcpectlq 13
      |||| | : | | :
Db      251 PDVQCGDCTLK 262

RESULT 30
FNHU
fibronectin precursor - human

```


N:Alternate names: fibronectin splice form ED-A
C:Species: Homo sapiens (man)
C:Date: 27-Nov-1985 #sequence.revision 31-Mar-1993 #text.change 22-Jun-1999
C:Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22
R:Dean, D.C.; Bowlus, C.L.; Bourgeois, S.
Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987
A:Title: Cloning and analysis of the promoter region of the human fibronectin gene.
A:Reference number: A26460; MUID:87175578
A:Accession: A26460
A:Molecule type: DNA
A:Residues: 1-49 <DEA>
A:Cross-references: GB:M15801; NID:g182686; PIDN:AAA53376.1; PID:g553293
R:Oldberg, A.; Ruoslahti, E.
J. Biol. Chem. 261, 2113-2116, 1986
A:Title: Evolution of the fibronectin gene.
A:Reference number: A26284; MUID:86111901
A:Accession: A26284
A:Molecule type: DNA
A:Residues: 1447-1540 <OLD>
A:Cross-references: GB:M12549; NID:g182688
A:Note: the authors translated the codon TTC for residue 1494 as Glu
R:Paolella, G.; Henochliffe, C.; Sebastio, G.; Baralle, F.E.
Nucleic Acids Res. 16, 3545-3557, 1988
A:Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B
A:Reference number: S00848; MUID:88233940
A:Accession: S03917
A:Molecule type: DNA
A:Residues: 1594-1767, 'V', 1769-1783 <PAO>
A:Cross-references: EMBL:X07718; NID:g31402
A:Note: the authors translated the codon AAC for residue 1631 as Asp
R:Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.
FEBS Lett. 207, 287-291, 1986
A:Title: Donor and acceptor splice signals within an exon of the human fibronectin gene.
A:Reference number: A24854; MUID:87030929
A:Accession: A24854
A:Molecule type: DNA
A:Residues: 1992-2147 <VIB>
A:Cross-references: GB:X04530; NID:g31436
R:Gutman, A.; Yamada, K.M.; Kornblihtt, A.
FEBS Lett. 207, 145-148, 1986
A:Title: Human fibronectin is synthesized as a pre-propolypeptide.
A:Reference number: A24476; MUID:87030890
A:Accession: A24476
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-14, 'Q', 16-38 <GUT>
R:Kornblihtt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.
EMBO J. 4, 1755-1759, 1985
A:Title: Primary structure of human fibronectin: differential splicing may generate at l
A:Reference number: A91008; MUID:85284965
A:Accession: A91008
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 32-1344, 1346-2080; 2112-2386 <KOR>
A:Cross-references: GB:X02761
R:Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Nucleic Acids Res. 12, 5853-5868, 1984
A:Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptid
A:Reference number: A93529; MUID:84272258
A:Accession: A93529
A:Molecule type: mRNA
A:Residues: 973-2080; 2112-2386 <K02>
A:Cross-references: GB:X00739
R:Oldberg, A.; Linney, E.; Ruoslahti, E.
J. Biol. Chem. 258, 10193-10196, 1983
A:Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell a
A:Reference number: A21011; MUID:83290929
A:Accession: A21011
A:Molecule type: mRNA
A:Residues: 1434-1537 <OL2>
A:Cross-references: GB:X00055; NID:g182680; PIDN:AAA52459.1; PID:g182683
R:Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
Biochemistry 24, 2698-2704, 1985

A:Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with
A:Reference number: A90495; MUID:85280409
A:Accession: A90495
A:Molecule type: mRNA
A:Residues: 1594-2386 <BER>
A:Cross-references: GB:M10905; NID:g182696; PIDN:AAA52462.1; PID:g182697
R:Umezawa, K.; Kornblihtt, A.R.; Baralle, F.E.
FEBS Lett. 186, 31-34, 1985
A:Title: Isolation and characterization of cDNA clones for human liver fibronectin.
A:Reference number: A22245; MUID:85231203
A:Accession: A22245
A:Molecule type: mRNA
A:Residues: 1975-1991; 2017-2039 <UM2>
A:Cross-references: GB:M27590
R:Seikiguchi, K.; Klos, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A:Title: Human liver fibronectin complementary DNAs: identification of two different
A:Reference number: 152394; MUID:87026578
A:Accession: 152373
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1978-1990, 2016-2018, 'N', 2020-2081, 2113-2127 <SEK>
A:Cross-references: GB:M14060; NID:g182701; PIDN:AAA52464.1; PID:g182704
R:Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A:Title: Isolation and characterization of cDNA clones for human and bovine fibronect
A:Reference number: A21165; MUID:83221567
A:Accession: A21165
A:Molecule type: mRNA
A:Residues: 2291-2386 <KO3>
A:Cross-references: GB:X00799; NID:g182681; PIDN:AAA52460.1; PID:g182684
R:Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 258, 12670-12674, 1983
A:Title: Primary structure of human plasma fibronectin.
A:Reference number: A92398; MUID:84032463
A:Accession: A92398
A:Molecule type: protein
A:Residues: 32-47, 'C', 49-51, 'S', 53-72, 'A', 74-290 <GAR1>
R:Garcia-Pardo, A.; Gold, L.I.
Arch. Biochem. Biophys. 304, 181-188, 1993
A:Title: Further characterization of the binding of fibronectin to gelatin reveals th
A:Reference number: S34791; MUID:93312001
A:Accession: S34791
A:Molecule type: protein
A:Residues: 291-300; 551-560 <GAR2>
R:Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
Thromb. Res. 43, 469-477, 1986
A:Title: Two plasma fibronectin fragments with different gelatin-binding properties.
A:Reference number: A60904; MUID:87019725
A:Accession: A60904
A:Molecule type: protein
A:Residues: 293-301 <GRI>
R:Calaycay, J.; Pande, H.; Lee, T.; Borsi, L.; Siri, A.; Shively, J.E.; Zardi, L.
J. Biol. Chem. 260, 12136-12141, 1985
A:Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human
A:Reference number: A23901; MUID:86008277
A:Accession: A23901
A:Molecule type: protein
A:Residues: 616-677, 'Q', 679-703, 'PT' <CAL>
R:Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
J. Biol. Chem. 257, 9593-9597, 1982
A:Title: The cell attachment domain of fibronectin. Determination of the primary stru
A:Reference number: A92386; MUID:82265604
A:Accession: A92386
A:Molecule type: protein
A:Residues: 1441-1548 <PIE>
A:Note: residues 1524-1527 are responsible for the cell-binding activity
R:Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
Biochem. J. 241, 923-928, 1987

F: 905-987/Domain: fibronectin type III repeat homology <FN3D>
F: 995-1076/Domain: fibronectin type III repeat homology <FN3E>
F: 1085-1164/Domain: fibronectin type III repeat homology <FN3F>
F: 1172-1257/Domain: fibronectin type III repeat homology <FN3G>
F: 1265-1348/Domain: fibronectin type III repeat homology <FN3H>
F: 1356-1439/Domain: fibronectin type III repeat homology <FN3I>
F: 1447-1529/Domain: fibronectin type III repeat homology <FN3J>
F: 1537-1619/Domain: fibronectin type III repeat homology <FN3K>
F: 1614-1616/Region: cell attachment (R-G-D) motif
F: 1621-1713/Domain: fibronectin type III repeat homology <FN3L>
F: 1731-1803/Domain: fibronectin type III repeat homology <FN3M>
F: 1803-1894/Domain: fibronectin type III repeat homology <FN3N>
F: 1901-1984/Domain: fibronectin type III repeat homology <FN3O>
F: 1992-2074/Domain: fibronectin type III repeat homology <FN3P>
F: 2181-2193/Region: cell attachment (R-G-D) motif
F: 2193-2273/Domain: fibronectin type III repeat homology <FN3Q>
F: 2296-2335/Domain: fibronectin type I repeat homology <1F10>
F: 2341-2378/Domain: fibronectin type I repeat homology <1F11>
F: 2385-2420/Domain: fibronectin type I repeat homology <1F12>
F: 53-79,77-88,98-126,124-136,142-170,168-180,187-216,214-226,232-368,2366-2378,2385-2411,2409-2420/Disulfide bonds: status predicted
F: 2458/Disulfide bonds: interchain (to 2462) #status predicted
F: 2462/Disulfide bonds: interchain (to 2458) #status predicted

Query Match 33.7%; Score 49.5; DB 2; Length 2477;
Best Local Similarity 42.9%; Pred. NO. 93;
Matches 9; Conservative 5; Mismatches 6; Indels 1

Qy 2 p d v q d c p e c t l g e n p f f s q p g 22
 | : : | | | : : | : | : | :

Dd 2098 P E I L D V P S - T V Q K T P F V T N P G 2117

RESULT 32

S25690
 hupJ protein - Rhodobacter capsulatus
 C:Species: Rhodobacter capsulatus
 C:Date: 04-Dec-1992 #sequence_revision 13-Mar-1997 #text_change 11-Jun-1999
 C:Accession: S32946; S25690; A38532
 R:Colbeau, A.; Richaud, P.; Toussaint, B.; Caballero, F.J.; Elster, C.; Delphin, C.; Smi
 Mol. Microbiol. 8, 15-29, 1993
 A:Title: Organization of the genes necessary for hydrogenase expression in Rhodobacter c
 A:Reference number: S32941; MUID:93268090
 A:Accession: S32946
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-278 <COL>
 A:Cross-references: EMBL:Z15089; NID:g313868; PIDN:CAA78802.1; PID:g581495
 A:Note: the authors translated the initiation codon TTG for residue 1 as Leu
 A:Note: the authors translated the codon ATA for residue 34 as B
 R:Toussaint, B.
 submitted to the EMBL Data Library, September 1992
 A:Reference number: S25686
 A:Accession: S25690
 A:Molecule type: DNA
 A:Residues: 26-278 <TOU>
 A:Cross-references: EMBL:Z15089
 R:Xu, H.W.; Wall, J.D.
 J. Bacteriol. 173, 2401-2405, 1991
 A:Title: Clustering of genes necessary for hydrogen oxidation in Rhodobacter capsulatus.
 A:Reference number: A38532; MUID:91177833
 A:Accession: A38532
 A:Molecule type: DNA
 A:Residues: 172-250, 'PGPRRWNRGGG' <XUA>
 A:Cross-references: GB:M55089; NID:gl51949; PIDN:AAA72923.1; PID:gl51950
 C:Genetics:
 A:Gene: hupJ
 A:Start codon: TTG
 A:Note: part of an operon containing 18 genes involved in hydrogenase activity and expre
 C:Superfamily: Rhodobacter hupJ protein; rubredoxin homology
 C:Keywords: electron transfer; iron
 F:25-71/Domain: rubredoxin homology <RUB>

F;28,31,61,64/Binding site: iron (Cys) #status predicted

Query Match 33.3%; Score 49; DB 1; Length 278;
Best Local Similarity 40.7%; Pred. NO. 13;
Matches 11; Conservative 2; Mismatches 12; Indels 2; Gaps 1;

Qy 2 pdvqdcpectlqenpff--sqpgapil 26
 |: ||| : | | |||||
 Db 56 PEDWHCPNCDAPKAQFIVQSDPGAPAL 82

RESULTS 33

I50992
 gonadotropin alpha-subunit - Morone saxatilis
 C:Species: Morone saxatilis
 C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 17-Mar-1999
 C:Accession: I50992
 E:Hassin, S.; Elizur, A.; Zohar, Y.
 J. Mol. Endocrinol. 15, 23-35, 1995
 A:Title: Molecular cloning and sequence analysis of striped bass (*Morone saxatilis*) g
 A:Reference number: I50992; MID:96020549
 A:Accession: I50992
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-117 <HAS>
 A:Cross-references: GB:IJ35071; NID:g522304; PID:g598254
 C:Superfamily: glycoprotein hormones alpha chain

Query Match 32.7%; Score 48; DB 2; Length 117;
Best Local Similarity 57.1%; Pred. No. 7.5;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 7 cpeclqenpfqs 20
| | | | | : | | :
Db 34 CEECTLRKNSVFSR 47

RESULT 34

B57062
SRB9 protein - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein YDR443c; SCAL protein
C:Species: *Saccharomyces cerevisiae*
C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 07-May-1999
C:Accession: B57062; S48538; S69722; S70382; S59987
R:Hangartner, C. J.; Thompson, C. M.; Zhang, J.; Chao, D. M.; Liao, S. M.; Koleske, A. J.;
Genes Dev. 9, 897-910, 1995
A:Title: Association of an activator with an RNA polymerase II holoenzyme.
A:Reference number: A57062; MUID:95293223
A:Accession: B57062
A:Molecule type: DNA
A:Residues: 1-1420 <HEN>
A:Cross-references: GB:U23812; NID:g780740; PID:g780741
R:Yuryev, A.; Corden, J.L.
Submitted to the EMBL Data Library, April 1994
A:Description: Suppression of Lethal Substitution Mutations in the C-terminal Domain
A:Reference number: S48538
A:Accession: S48538
A:Molecule type: DNA
A:Residues: 1-37, 'E', 39-811, 'V', 813-858, 'S', 860-876, 'GE', 879-886, 'P', 888-1283, 'S', 1283-1289, 'E', 1290-1291, 'E', 1292-1293, 'E', 1294-1295, 'E', 1296-1297, 'E', 1298-1299, 'E', 1300-1301, 'E', 1302-1303, 'E', 1304-1305, 'E', 1306-1307, 'E', 1308-1309, 'E', 1310-1311, 'E', 1312-1313, 'E', 1314-1315, 'E', 1316-1317, 'E', 1318-1319, 'E', 1320-1321, 'E', 1322-1323, 'E', 1324-1325, 'E', 1326-1327, 'E', 1328-1329, 'E', 1330-1331, 'E', 1332-1333, 'E', 1334-1335, 'E', 1336-1337, 'E', 1338-1339, 'E', 1340-1341, 'E', 1342-1343, 'E', 1344-1345, 'E', 1346-1347, 'E', 1348-1349, 'E', 1350-1351, 'E', 1352-1353, 'E', 1354-1355, 'E', 1356-1357, 'E', 1358-1359, 'E', 1360-1361, 'E', 1362-1363, 'E', 1364-1365, 'E', 1366-1367, 'E', 1368-1369, 'E', 1370-1371, 'E', 1372-1373, 'E', 1374-1375, 'E', 1376-1377, 'E', 1378-1379, 'E', 1380-1381, 'E', 1382-1383, 'E', 1384-1385, 'E', 1386-1387, 'E', 1388-1389, 'E', 1390-1391, 'E', 1392-1393, 'E', 1394-1395, 'E', 1396-1397, 'E', 1398-1399, 'E', 1400-1401, 'E', 1402-1403, 'E', 1404-1405, 'E', 1406-1407, 'E', 1408-1409, 'E', 1410-1411, 'E', 1412-1413, 'E', 1414-1415, 'E', 1416-1417, 'E', 1418-1419, 'E', 1420-1421, 'E', 1422-1423, 'E', 1424-1425, 'E', 1426-1427, 'E', 1428-1429, 'E', 1430-1431, 'E', 1432-1433, 'E', 1434-1435, 'E', 1436-1437, 'E', 1438-1439, 'E', 1440-1441, 'E', 1442-1443, 'E', 1444-1445, 'E', 1446-1447, 'E', 1448-1449, 'E', 1450-1451, 'E', 1452-1453, 'E', 1454-1455, 'E', 1456-1457, 'E', 1458-1459, 'E', 1460-1461, 'E', 1462-1463, 'E', 1464-1465, 'E', 1466-1467, 'E', 1468-1469, 'E', 1470-1471, 'E', 1472-1473, 'E', 1474-1475, 'E', 1476-1477, 'E', 1478-1479, 'E', 1480-1481, 'E', 1482-1483, 'E', 1484-1485, 'E', 1486-1487, 'E', 1488-1489, 'E', 1490-1491, 'E', 1492-1493, 'E', 1494-1495, 'E', 1496-1497, 'E', 1498-1499, 'E', 1500-1501, 'E', 1502-1503, 'E', 1504-1505, 'E', 1506-1507, 'E', 1508-1509, 'E', 1510-1511, 'E', 1512-1513, 'E', 1514-1515, 'E', 1516-1517, 'E', 1518-1519, 'E', 1520-1521, 'E', 1522-1523, 'E', 1524-1525, 'E', 1526-1527, 'E', 1528-1529, 'E', 1530-1531, 'E', 1532-1533, 'E', 1534-1535, 'E', 1536-1537, 'E', 1538-1539, 'E', 1540-1541, 'E', 1542-1543, 'E', 1544-1545, 'E', 1546-1547, 'E', 1548-1549, 'E', 1550-1551, 'E', 1552-1553, 'E', 1554-1555, 'E', 1556-1557, 'E', 1558-1559, 'E', 1560-1561, 'E', 1562-1563, 'E', 1564-1565, 'E', 1566-1567, 'E', 1568-1569, 'E', 1570-1571, 'E', 1572-1573, 'E', 1574-1575, 'E', 1576-1577, 'E', 1578-1579, 'E', 1580-1581, 'E', 1582-1583, 'E', 1584-1585, 'E', 1586-1587, 'E', 1588-1589, 'E', 1590-1591, 'E', 1592-1593, 'E', 1594-1595, 'E', 1596-1597, 'E', 1598-1599, 'E', 1600-1601, 'E', 1602-1603, 'E', 1604-1605, 'E', 1606-1607, 'E', 1608-1609, 'E', 1610-1611, 'E', 1612-1613, 'E', 1614-1615, 'E', 1616-1617, 'E', 1618-1619, 'E', 1620-1621, 'E', 1622-1623, 'E', 1624-1625, 'E', 1626-1627, 'E', 1628-1629, 'E', 1630-1631, 'E', 1632-1633, 'E', 1634-1635, 'E', 1636-1637, 'E', 1638-1639, 'E', 1640-1641, 'E', 1642-1643, 'E', 1644-1645, 'E', 1646-1647, 'E', 1648-1649, 'E', 1650-1651, 'E', 1652-1653, 'E', 1654-1655, 'E', 1656-1657, 'E', 1658-1659, 'E', 1660-1661, 'E', 1662-1663, 'E', 1664-1665, 'E', 1666-1667, 'E', 1668-1669, 'E', 1670-1671, 'E', 1672-1673, 'E', 1674-1675, 'E', 1676-1677, 'E', 1678-1679, 'E', 1680-1681, 'E', 1682-1683, 'E', 1684-1685, 'E', 1686-1687, 'E', 1688-1689, 'E', 1690-1691, 'E', 1692-1693, 'E', 1694-1695, 'E', 1696-1697, 'E', 1698-1699, 'E', 1700-1701, 'E', 1702-1703, 'E', 1704-1705, 'E', 1706-1707, 'E', 1708-1709, 'E', 1710-1711, 'E', 1712-1713, 'E', 1714-1715, 'E', 1716-1717, 'E', 1718-1719, 'E', 1720-1721, 'E', 1722-1723, 'E', 1724-1725, 'E', 1726-1727, 'E', 1728-1729, 'E', 1730-1731, 'E', 1732-1733, 'E', 1734-1735, 'E', 1736-1737, 'E', 1738-1739, 'E', 1740-1741, 'E', 1742-1743, 'E', 1744-1745, 'E', 1746-1747, 'E', 1748-1749, 'E', 1750-1751, 'E', 1752-1753, 'E', 1754-1755, 'E', 1756-1757, 'E', 1758-1759, 'E', 1760-1761, 'E', 1762-1763, 'E', 1764-1765, 'E', 1766-1767, 'E', 1768-1769, 'E', 1770-1771, 'E', 1772-1773, 'E', 1774-1775, 'E', 1776-1777, 'E', 1778-1779, 'E', 1780-1781, 'E', 1782-1783, 'E', 1784-1785, 'E', 1786-1787, 'E', 1788-1789, 'E', 1790-1791, 'E', 1792-1793, 'E', 1794-1795, 'E', 1796-1797, 'E', 1798-1799, 'E', 1800-1801,

A:Reference number: S70382; MUID:96363903

A:Accession: S70382

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-37, 'E', 39-811, 'Y', 813-858, 'S', 860-876, 'GE', 879-886, 'P', 888-1283, 'S', 1285-1

A:Cross-references: EMBL:U09176; NID:g495489; PID:g495490

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994

C:Genetics:

A:Gene: SGD:SSN2; SRB9; SCAL

A:Cross-references: MIPS:YDR443c; SGD:S0002851

A:Map position: 4R

Query Match 32.7%; Score 48; DB 2; Length 1420;

Best Local Similarity 45.5%; Pred. No. 88;

Matches 10; Conservative 3; Mismatches 7; Indels 2; Gaps 1;

QY 6 dcp--ectlqenpfsgpapi 25

||| ||| ||| ||| |||

Db 576 DIPIDEMTLPTSPLYMPDGAPL 597

RESULT 35

D64392

C:Species: Methanococcus jannaschii

C:Accession: D64392

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 11-Jun-1999

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,

; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;

ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A:Reference number: A64300; MUID:96337999

A:Accession: D64392

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-55 <BUL>

A:Cross-references: GB:U67520; GB:L77117; NID:gl591447; PIDN:AB98734.1; PID:gl591453; T

C:Genetics:

A:Map position: REV668868-668701

A:Start codon: GTG

C:Superfamily: rubredoxin; rubredoxin homology

F:6-51/Domain: rubredoxin homology <RUB>

Query Match 32.0%; Score 47; DB 2; Length 55;

Best Local Similarity 50.0%; Pred. No. 4,9;

Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 pdvqdcpectlqenpf 17

||| ||| ||| ||| |||

Db 36 PDTRPCPOGLGNKF 51

RESULT 36

T26821

C:Species: Caenorhabditis elegans

C:Accession: T26821

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

R:Steward, C.

submitted to the EMBL Data Library, October 1998

A:Reference number: Z20271

A:Accession: T26821

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-114 <WIL>

A:Cross-references: EMBL:AL032618; PIDN:CAA21486.1; CESP:Y42A5A.3

A:Experimental source: clone Y42A5A

C:Genetics:

A:Gene: CESP:Y42A5A.3

A:Introns: 54/3; 79/2

Query Match 32.0%; Score 47; DB 2; Length 114;

Best Local Similarity 50.0%; Pred. No. 10;

Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 8 pectlqenpfsgpapi 25

||| ||| ||| ||| |||

Db 65 PIVTRQNTFIDODGVVP 82

RESULT 37

RNB29

DNA-directed RNA polymerase (EC 2.7.7.6) II chain RPB9 - yeast (Saccharomyces cerevisiae)

N:Alternate names: DNA-directed RNA polymerase 14.5K chain; protein G3271; protein 1G

C:Species: Saccharomyces cerevisiae

C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 11-Jun-1999

C:Accession: A41016; S19043; S64077

R:Woychik, N.A.; Lane, W.S.; Young, R.A.

J. Biol. Chem. 266, 19053-19055, 1991

A:Title: Yeast RNA polymerase II subunit RPB9 is essential for growth at temperature

A:Reference number: A41016; MUID:92011681

A:Accession: A41016

A:Molecule type: DNA

A:Residues: 1-122 <WOY1>

A:Cross-references: GB:M73060; NID:gl72469; PIDN:AAA34997.1; PID:gl72470

A:Accession: S19043

A:Molecule type: protein

A:Residues: 46-62 <WOY2>

R:Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.

submitted to the Protein Sequence Database, May 1996

A:Reference number: S64071

A:Accession: S64077

A:Molecule type: DNA

A:Residues: 1-122 <RIE>

A:Cross-references: EMBL:Z72592; NID:gl322580; PIDN:CAA96774.1; PID:e243958; PID:gl32

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:RPB9

A:Cross-references: SGD:S0003038; MIPS:YGL070c

A:Map position: 7L

C:Superfamily: DNA-directed RNA polymerase II chain RPB9

C:Keywords: nucleotidyltransferase; nucleus; transcription; zinc finger

F:7-32/Region: zinc finger CCCC motif

F:75-106/Region: zinc finger CCCC motif

A:Accession: A43821
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-260 <ERS>
A:Cross-references: GB:X55318; NID:g57878; PIDN:CAA39026.1; PID:g57879
A:Note: sequence extracted from NCBI backbone (NCBIN:83339, NCBI:P:83341)
R:Breier, G.; Dressler, G.R.; Gruss, P.
EMBO J. 7, 1329-1336, 1988
A:Title: Primary structure and developmental expression pattern of Hox 3.1, a member of
A:Reference number: S00548; MUID:98312579
A:Accession: S01194
A:Molecule type: DNA
A:Residues: 187-260 <BRE>
A:Cross-references: EMBL:X07647; NID:g51408; PIDN:CAA30487.1; PID:es290; PID:gl333933
R:Anguilewitsch, A.; Bieberich, C.; Bogardt, L.; Shashikant, C.; Ruddle, F.H.
Proc. Natl. Acad. Sci. U.S.A. 87, 6428-6432, 1990
A:Title: Structural analysis of the Hox-3.1 transcription unit and the Hox-3.2-Hox-3.1
A:Reference number: A36023; MUID:90349629

C;Superfamily: Enterobacter ribonuclease


```
Query Match      31.3%; Score 46; DB 2; Length 227;
Best Local Similarity 52.4%; Pred. No. 28;
Matches 11; Conservative 1; Mismatches 1; Indels 8; Gaps 2;

Qy 6 dceptlqenffsgpapl 26
||| ||| |||
Db 207 DCP-----NPF-QPGSPYL 219

RESULT 43
P3VXP5
3a protein - peanut stunt virus (strain J)
C:Species: peanut stunt virus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C:Accession: A40786; JQ0504
R:Karasawa, A.; Nakaho, K.; Kakutani, T.; Minobe, Y.; Ehara, Y.
Virology 185, 464-467, 1991
A:Title: Nucleotide sequence of RNA 3 of peanut stunt cucumovirus.
A:Reference number: A40786; MUID:92024111
A:Accession: A40786
A:Molecule type: genomic RNA
A:Residues: 1-288 <KAR>
A:Cross-references: GB:D00668; NID:g222421; PIDN:BAA00571.1; PID:d1001027; PID:g222422
C:Genetics:
A:Map position: segment 3
C:Superfamily: bromo mosaic virus 3a protein
C:Keywords: nonstructural protein

Query Match      31.3%; Score 46; DB 1; Length 288;
Best Local Similarity 38.9%; Pred. No. 35;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 9 ectlqenffsgpapl 26
: | : | | | | |
Db 224 QLNWQSSPLFQLPGGPIM 241

RESULT 44
S63536
aminomethyltransferase (EC 2.1.2.10) gdcSPB precursor - Flaveria pringlei
N:Alternate names: glycine cleavage system protein PB
C:Species: Flaveria pringlei
C:Date: 13-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 22-Jun-1999
C:Accession: S63536
R:Bauwe, H.; Chu, C.-C.; Kopriva, S.; Nan, Q.
Eur. J. Biochem. 234, 116-124, 1995
A:Title: Structure and expression analysis of the gdcSPA and gdcSPB genes encoding two P
A:Reference number: S63536; MUID:96096729
A:Accession: S63536
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1034 <BAU>
A:Cross-references: EMBL:Z54239; NID:g1000488; PID:g1000489
C:Genetics:
A:Genome: nuclear
A:Introns: 401/3; 509/3; 543/3; 608/3; 650/3; 712/3; 750/3; 801/3; 864/3; 890/3;
C:Keywords: mitochondrion; phosphoprotein; pyridoxal phosphate; transferase
F:1-63/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F:64-1034/Product: aminomethyltransferase gdcSPB #status predicted <MAT>
F:770/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match      31.0%; Score 45.5; DB 2; Length 1034;
Best Local Similarity 40.9%; Pred. No. 1.4e+02;
Matches 9; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

Qy 1 apdvqdc-pectlqenffsq 21
||:|||| |::| |::|
Db 518 APEVQDAIPSGLVRETPYLTHP 539

RESULT 45
S63535
aminomethyltransferase (EC 2.1.2.10) gdcSPA precursor - Flaveria pringlei
N:Alternate names: glycine cleavage system protein PA
C:Species: Flaveria pringlei
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 22-Jun-1999
C:Accession: S63535
R:Bauwe, H.; Chu, C.-C.; Kopriva, S.; Nan, Q.
Eur. J. Biochem. 234, 116-124, 1995
A:Title: Structure and expression analysis of the gdcSPA and gdcSPB genes encoding tw
A:Reference number: S63535; MUID:96096729
A:Accession: S63535
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1037 <BAU>
A:Cross-references: EMBL:Z36879; NID:g608711; PID:g608712
C:Genetics:
A:Genome: nuclear
A:Introns: 404/3; 486/3; 512/3; 546/2; 611/3; 653/3; 693/1; 715/3; 753/3; 804/3; 867/
C:Keywords: mitochondrion; phosphoprotein; pyridoxal phosphate; transferase
F:1-66/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F:67-1037/Product: aminomethyltransferase gdcSPA #status predicted <MAT>
F:773/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match      31.0%; Score 45.5; DB 2; Length 1037;
Best Local Similarity 40.9%; Pred. No. 1.4e+02;
Matches 9; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

Qy 1 apdvqdc-pectlqenffsq 21
||:|||| |::| |::|
Db 521 APEVQDAIPSGLVRETPYLTHP 542

RESULT 46
S40216
P protein - Flaveria pringlei
C:Species: Flaveria pringlei
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 22-Jun-1999
C:Accession: S40216
R:Kopriva, S.; Bauwe, H.
submitted to the EMBL Data Library, September 1993
A:Description: P-protein of the glycine decarboxylase multienzyme complex from Flaver
A:Reference number: S40211
A:Accession: S40216
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1037 <KOP>
A:Cross-references: EMBL:Z25857; NID:g438002; PID:g438003
C:Keywords: phosphoprotein; pyridoxal phosphate
F:773/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match      31.0%; Score 45.5; DB 2; Length 1037;
Best Local Similarity 40.9%; Pred. No. 1.4e+02;
Matches 9; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

Qy 1 apdvqdc-pectlqenffsq 21
||:|||| |::| |::|
Db 521 APEVQDAIPSGLVRETPYLTHP 542

RESULT 47
W5WLR1
E5 protein - rhesus papillomavirus (type 1)
C:Species: rhesus papillomavirus
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 27-Jan-1995
C:Accession: F38503
R:Ostrow, R.S.; LaBresh, K.V.; Faras, A.J.
Virology 181, 424-429, 1991
A:Title: Characterization of the complete RHPV 1 genomic sequence and an integration
A:Reference number: A38503; MUID:91135018
```


A:Accession: F38503
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-157 <OST>
A:Cross-references: EMBL:M37717
C:Superfamily: rhesus papillomavirus E5 protein
C:Keywords: early protein

Query Match 30.6%; Score 45; DB 1; Length 157;
Best Local Similarity 53.8%; Pred. No. 26;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 dvqdcpectlqen 15
| | | | | | | | | |
Db 104 DTPACPCGLQQN 116

RESULT 48

hypothetical protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C:Accession: G71427
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Hempel, S.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Gied
avanagh, T.; Kempel, S.; Kottler, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; And
C.; Chalwatis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A:Reference number: A71400; MUID:98121113
A:Accession: G71427
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-164 <BEV>
A:Cross-references: GB:T97340; NID:g2244950; PID:e326975; PID:g2244972
C:Genetics:
A:Map position: 4COP9-4G3845

Query Match 30.6%; Score 45; DB 2; Length 164;
Best Local Similarity 38.1%; Pred. No. 28;
Matches 8; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 4 vqdcpectlqenpfsgpgap 24
| | | | | | | | | |
Db 28 IADCTMCTCDNCPQNPSP 48

RESULT 49

G75403
DNA topoisomerase I - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: G75403
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250
A:Accession: G75403
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1021 <WHI>
A:Cross-references: GB:AE001983; NID:g6459123; PID:AAF10943.1; PID:g645912
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1374
A:Map position: 1

Query Match 30.6%; Score 45; DB 2; Length 1021;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 11 tlqenpfsgpgapil 26
| | | | | | | | | |
Db 320 TAEKPFSTRPPAPFI 335

RESULT 50

T09361
hypothetical protein F23K16.80 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 24-Nov-1999
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16652
A:Accession: T09361
A:Molecule type: DNA
A:Residues: 1-1553 <BEV>
A:Cross-references: EMBL:AL078620; GSPDB:GN00062; ATSP:F23K16.80
A:Experimental source: cultivar Columbia; BAC clone F23K16
C:Genetics:
A:Gene: ATSP:F23K16.80
A:Map position: 4
A:Introns: 8/3; 65/3; 132/3; 319/3; 428/2; 566/3; 721/3; 760/3; 801/2; 1044/3; 1129/3
C:Superfamily: Arabidopsis thaliana hypothetical protein F23K16.80

Query Match 30.6%; Score 45; DB 2; Length 1553;
Best Local Similarity 37.0%; Pred. No. 2.5e+02;
Matches 10; Conservative 6; Mismatches 7; Indels 4; Gaps 2;

QY 2 pdvqdcpectlqenpf--sqpgapil 26
| | | | | | | | | |
Db 221 PDISSVPR--QKQNPSTSGSQGLPLM 245

Search completed: July 14, 2000, 09:33:32
Job time: 2265 sec


```
> 0 <
0! 10 IntelliGenetics
> 0 <

Quest - Quick User-directed Expression Search Tool
Release 5.4

    -- Outline of search "beta-58" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "new.key":
beta-58 (AA) ID beta-58 AA preliminary pattern
1 followed by
2 vctyrdf
2 r or h or k
2 yrtv
2 e or r or h or k
2 ipgcp
2 l or r or h or k
2 hvapyfsypva
2

Selected data banks and files:

Data bank : A-Geneseq 35.2, all entries
Data bank : Issued_AA , all entries
Data bank : Pending_AA , all entries

    -- Output Parameters --

Format Options:
Nucleic acid code matching Exact Indirect file No
Find non-matching hits only No Sequence or key file No
Report key used Yes List of hits Yes
Note position of hit Yes Hit display Yes
Display full annotations Yes Name and annotations Yes
Sequence context 10

    -- Run Parameters --

Run mode Batch
Time to start comparison now
Notify at end of run No

No hits found.

    -- Search Statistics --

Times: CPU 00:20:42.11 Total Elapsed 00:35:02.00

Number of sequences searched: 1068391
Number of sequence hits: 0
Number of separate matches: 0
Number of sequence hits saved: 0
```



```
> O <
OI JO Intelligenetics
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

-- Outline of search "beta-63" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "new.key":
beta-63 (AA) ID beta-63 AA preliminary pattern
1 followed by
2 vctyrdff
2 i or r or h or k
2 yrtv
2 r or h or k
2 ipgcp
2 l or r or h or k
2 hvapyfsypva

Selected data banks and files:

Data bank : A-Geneseq 35.2, all entries
Data bank : Issued_AA , all entries
Data bank : Pending_AA , all entries

-- Output Parameters --

Format Options:                               File Options:
Nucleic acid code matching      Exact          Indirect file      No
Find non-matching hits only    No          Sequence or key file   No
Report key used                Yes          List of hits           Yes
Note position of hit           Yes          Hit display            Yes
Display full annotations        Yes          Name and annotations    Yes
Sequence context                10

-- Run Parameters --

Run mode                          Batch
Time to start comparison         now
Notify at end of run             No

No hits found.

-- Search Statistics --

Times:                               CPU                               Total Elapsed
                                00:20:54.05                               00:35:36.00

Number of sequences searched:      1068391
Number of sequence hits:           0
Number of separate matches:        0
Number of sequence hits saved:     0
```



```
> O <
O| IO  Intelligence
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

-- Outline of search "beta-69" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "new.key":
beta-69 (AA) ID beta-69 AA preliminary pattern
1 followed by
2 vctyrdf
2 i or r or h or k
2 yrtv
2 e or r or h or k
2 ipgcp
2 r or h or k
2 hvapyfsypva

Selected data banks and files:

Data bank : A-Geneseq 35.2, all entries
Data bank : Issued_AA , all entries
Data bank : Pending_AA , all entries

-- Output Parameters --

Format Options:
Nucleic acid code matching Exact No Indirect file NO
Find non-matching hits only Yes No Sequence or key file NO
Report key used Yes Yes List of hits Yes
Note position of hit Yes Hit display Yes
Display full annotations Yes Name and annotations Yes
Sequence context 10

-- Run Parameters --

Run mode Batch
Time to start comparison now
Notify at end of run No

No hits found.

-- Search Statistics --

Times: CPU 00:20:28.03 Total Elapsed 00:35:28.00
Number of sequences searched: 1068391
Number of sequence hits: 0
Number of separate matches: 0
Number of sequence hits saved: 0
```


! FINDPATTERNS on pir:* allowing 0 mismatches

! 1 APDVODCPEC(R,K,H)L(Q,R,H,K)(E,R,H,K)N(P,R,H,K)(F,R,H,K)FS(Q,R,H,K)PGAPIL

pattern searched

Databases searched:

NBRF, Release 63.0, Released on 30Dec1999, Formatted on 7Mar2000

Total finds: 0

Total length: 58,629,743

Total sequences: 166,808

CPU time: 03:05.78

! FINDPATTERNS on swp:* allowing 0 mismatches

! 1 APDVQDCPEC(R,K,H)L(Q,R,H,K)(E,R,H,K)N(P,R,H,K)(F,R,H,K)FS(Q,R,H,K)PGAPIL

- pattern searched

Databases searched:

SWISS-PROT, Release 38.0, Released on 15Feb2000, Formatted on 16Feb2000

SPTREMBL, Release 12.0, Released on 1Nov1999, Formatted on 27Dec1999

Total finds: 0

Total length: 99,789,095

Total sequences: 309,735

CPU time: 05:08.54

! FINDPATTERNS on pir:* allowing 0 mismatches

! 1 APDVQDCPEC(T,R,K,H)L(R,H,K)(E,R,H,K)N(P,R,H,K)(F,R,H,K)FS(Q,R,H,K)PGAPIL

- pattern searched

Databases searched:

NBRF, Release 63.0, Released on 30Dec1999, Formatted on 7Mar2000

Total finds: 0
Total length: 58,629,743
Total sequences: 168,808
CPU time: 03:04.77

! FINDPATTERNS on swp:* allowing 0 mismatches

! 1 APDVQDCPEC(T,R,K,H)L(R,H,K)(E,R,H,K)N(P,R,H,K)(F,R,H,K)FS(Q,R,H,K)PGAPIL

- pattern searched

Databases searched:

SWISS-PROT, Release 38.0, Released on 15Feb2000, Formatted on 16Feb2000

SPTREMBL, Release 12.0, Released on 1Nov1999, Formatted on 27Dec1999

Total finds: 0

Total length: 99,789,095

Total sequences: 309,735

CPU time: 05:05.37

! FINDPATTERNS on pir:* allowing 0 mismatches

! 1 APDVQDCPEC(T,R,K,H)L(Q,R,H,K)(R,H,K)N(P,R,H,K)(F,R,H,K)FS(Q,R,H,K)PGAPIL

pattern searched

Databases searched:

NRKF, Release 63.0, Released on 30Dec1999, Formatted on 7Mar2000

Total finds: 0

Total length: 58,629,743

Total sequences: 168,808

CPU time: 03:07.38

! FINDPATTERNS on swp:* allowing 0 mismatches

! 1 APDVODCEC(T,R,K,H)L(Q,R,H,K)(R,H,K)N(P,R,H,K)(F,R,H,K)FS(O,R,H,K)PGAPIL

- pattern searched

Databases searched:

SWISS-PROT, Release 38.0, Released on 15Feb2000, Formatted on 16Feb2000
SPTREMBL, Release 12.0, Released on 1Nov1999, Formatted on 27Dec1999

Total finds: 0
Total length: 99,789,095
Total sequences: 309,735
CPU time: 05:09.26

! FINDPATTERNS on pir:* allowing 0 mismatches

! 1 APDVQDCPEC(T,R,K,H)L(Q,R,H,K)(E,R,H,K)N(R,H,K)(F,R,H,K)FS(Q,R,H,K)PGAPIL

- pattern searched

Databases searched:

NBRF, Release 63.0, Released on 30Dec1999, Formatted on 7Mar2000

Total finds: 0

Total length: 58,629,743

Total sequences: 188,808

CPU time: 03:05.77

! FINDPATTERNS on swp:* allowing 0 mismatches

! 1 APDVQDCPEC(T,R,K,H)L(Q,R,H,K)(E,R,H,K)N(R,H,K)(F,R,H,K)FS(Q,R,H,K)PGAPIL - *pattern searched*

Databases searched:

SWISS-PROT, Release 38.0, Released on 15Feb2000, Formatted on 16Feb2000

SPTREMBL, Release 12.0, Released on 1Nov1999, Formatted on 27Dec1999

Total finds: 0

Total length: 99,789,095

Total sequences: 309,735

CPU time: 05:10.25

! FINDPATTERNS on pir: * allowing 0 mismatches

! 1 APDVQDCPEC(T,R,K,H)L(Q,R,H,K){E,R,H,K}N(P,R,H,K)(R,H,K)FS(Q,R,H,K)PGADIL

pattern searched

Databases searched:

NBRF, Release 63.0, Released on 30Dec1999, Formatted on 7Mar2000

Total finds: 0

Total length: 58,629,743

Total sequences: 168,808

CPU time: 03:05.17

! FINDPATTERNS on swp:* allowing 0 mismatches

! 1 APDVQDCPEC(T,R,K,H)L(Q,R,H,K)(E,R,H,K)N(P,R,H,K)(R,H,K)FS(Q,R,H,K)PGAPIL

- pattern searched

Databases searched:

SWISS-PROT, Release 38.0, Released on 15Feb2000, Formatted on 16Feb2000

SPTREMBL, Release 12.0, Released on 1Nov1999, Formatted on 27Dec1999

Total finds: 0

Total length: 99,789,095

Total sequences: 309,735

CPU time: 05:08.08